From the Department of Clinical Neuroscience
Karolinska Institutet, Stockholm, Sweden

MULTIPLE SCLEROSIS: FROM GENETIC VARIANTS TO BIOMARKERS

Sahl Khalid Bedri

Karolinska Institutet

Stockholm 2019

Published by Karolinska Institutet.
Printed by E-print AB 2019

ISBN 978-91-7831-553-6

Multiple sclerosis: from genetic variants to biomarkers

THESIS FOR DOCTORAL DEGREE (Ph.D.)

By

Sahl Khalid Bedri

Public defence: Friday 27th of September, 2019 at 9:00, at the
Centre for Psychiatry Research, Norra Stationsgatan 69, 6th floor, 90-salen

Principal Supervisor:
Associate Professor Anna Glaser
Karolinska Institutet
Department of Clinical Neuroscience

Co-supervisor(s):
Professor Jan Hillert
Karolinska Institutet
Department of Clinical Neuroscience

Associate Professor Hans Grönlund
Karolinska Institutet
Department of Clinical Neuroscience

Dr. Wangko Lundström
Karolinska Institutet
Department of Clinical Neuroscience

Opponent:
Professor Laura Airas
University of Turku
Department of Clinical Medicine

Examination Board:
Professor Torbjörn Gräslund
Kungliga Tekniska Högskolan
Department of Protein science

Associate Professor Benedict Chambers
Karolinska Institutet
Department of Medicine, Huddinge

Professor Caroline Graff
Karolinska Institutet
Department of Neurobiology, Care Sciences and Society

To all multiple sclerosis patients who contributed samples for research .

ABSTRACT

Multiple sclerosis (MS) is a common chronic autoimmune and neurodegenerative disease of the central nervous system (CNS). MS is a debilitating disease that affects young adults, especially females. Why we develop MS? Is thought to be a consequence of our genes and the environment. In the past four years, the number of MS associated genetic variants have increased up to more than 200, however they explain only about 30% of its heritability.

The general aim of the research presented in this thesis is to explore genetic variants associated to MS and to employ our knowledge of the known associations to study their potentiality as biomarkers.

In paper I, we aimed to identify genetic variants that distinguish the relapsing relapsing remitting (RR)MS from the primary progressive (PP)MS courses using whole exome sequencing data. We report a number of common and rare variants that are associated to either course. Moreover, we identified enrichment of mutations of other progressive neurological disorders in PPMS patients.

In paper II we investigated the possibility of somatic mosaicism within the CNS resulting in sub-populations of cells involved in MS pathogenesis. We identified somatic genetic variants of the copy number variations (CNVs) type in the T cell receptor loci. These CNVs lead us to profile and compare the TCR repertoire in cells in the periphery and in the CNS.

A number of potent treatments have been introduced for treating MS patients and they succeeded in providing a better quality of life. In papers III and IV, based on the available information of MS associated genetic variants we studied the effects of two drugs, natalizumab and fingolimod, on the intra-individual profile of proteins within selected pathways with the ambition to identify biomarkers for MS treatment. We took a candidate gene approach in paper III and studied the treatment effects on soluble cytokine receptors and observed changes in plasma levels of sIL-7Rα, sIL-2Rα and sgp130. In paper IV we took a multiplex approach utilizing protein arrays and detected a decrease in plasma levels of nine proteins during natalizumab treatment. Furthermore, we validated and replicated the change for the most significant protein, PEBP1.

Hopefully the identified genetic variants and observed changes on the molecular level during treatment could pave the way to hypotheses generation in order to identify pathways affected by these variations and provide insight into the immunopathology of MS.

LIST OF SCIENTIFIC PAPERS

I. **Whole Exome Sequencing to Identify Genetic Variants Associated with Primary-Progressive Multiple Sclerosis**
Tojo James, Sahl Khalid Bedri, Paola Bronson, K.D. Nguyen, Karol Estrada, Aaron Day-Williams, Lars Alfredsson, Tomas Olsson, Anna Glaser, Jan Hillert, Ingrid Kockum
Manuscript

II. **Genomic comparison of immune cells in the periphery and the central nervous system in multiple sclerosis patients**
Sahl Khalid Bedri, Björn Evertsson, Mohsen Khademi, Tomas Olsson, Jan Hillert and Anna Glaser
Manuscript

III. **MS treatment effects on plasma cytokine receptor levels**
Sahl Khalid Bedri, Katharina Fink, Ali Manouchehrinia, Wangko Lundström, Ingrid Kockum, Tomas Olsson, Jan Hillert and Anna Glaser
Clinical Immunology 2018 Feb;187:15-25

IV. **Plasma protein profiling reveals candidate biomarkers for multiple sclerosis treatment**
Sahl Khalid Bedri, Ola B. Nilsson, Katharina Fink, Anna Månberg, Carl Hamsten, Burcu Ayoglu, Peter Nilsson, Tomas Olsson, Jan Hillert, Hans Grönlund, Anna Glaser
PLoS One. 2019;14(5):e0217208

CONTENTS

LIST OF ABBREVIATIONS

ADAs	anti-drug antibodies
ARR	annualized relapse rate
BBB	blood brain barrier
BCR	B cell receptor
CIS	clinically isolated syndrome
CNS	central nervous system
CNV	copy number variation
CSF	cerebrospinal fluid
DMF	dimethyl fumarate
DZ	dizygotic
EAE	experimental autoimmune encephalomyelitis
EBV	Epstein Barr virus
EDSS	expanded Disability Status Scale
GWAS	genome wide associations studies
HCNP-pp	hippocampal cholinergic neurostimulating peptide precursor protein
HLA	human leukocyte antigen
IFN-g	interferon gamma
IFNb	interferon beta
IL2RA	interleukin 2 receptor alpha
IL7RA	interleukin 7 receptor
IMSE	Immunomodulation and MS Epidemiology
IMSGC	international MS genetic consortium; IMSGC
JCV	John Cunningham virus
MAPK	mitogen activated protein kinase
MBP	myelin basic protein
MHC	major histocompatibility complex
MRI	magnetic resonance image

MS	multiple sclerosis
MZ	monozygotic
NAWM	normal appearing white matter
NfL	neurofilament light
Nrf2	nuclear factor (erythroid-derived 2)-like 2
OCBs	oligoclonal bands
PB	peripheral blood
PBC	population-based control
PBMC	peripheral blood mononuclear cell
PEBP1	phosphatidylethanolamine binding protein 1
PML	progressive multifocal leukoencephalopathy
PPMS	primary progressive MS
RIS	radiologically isolated syndrome
RKIP	RAF1 kinase inhibitor
RRMS	relapsing-remitting MS
sIL-7Rα	soluble interleukin-7 receptor alpha
sIL-2Rα	soluble interleukin-2 receptor alpha
sIL-6R	soluble interleukin-6 receptor
sgp130	soluble glycoprotein 130
S1PR1	sphinogosine-1 phosphate receptor
SNPs	single nucleotide polymorphisms
SPMS	secondary progressive MS
TCR	T cell receptor
TERT	telomerase reverse transcriptase
Tregs	regulatory T cells

1 INTRODUCTION

1.1 MULTIPLE SCLEROSIS

Multiple sclerosis (MS) is a chronic autoimmune and neurodegenerative disease of the central nervous system (CNS)[1]. After trauma MS is one of the most common causes for neurological disability in young adults. The description of MS can be followed back as early as to the 14th century in documents from the Vatican describing a Dutch nun, Saint Ludwina of Schiedam, who experienced episodes of leg weakness and visual disturbances interrupted by remission[2]. Five centuries later, it was the French neurologist Jean-Martin Charcot who connected both the neurological symptoms and pathological findings with the brain and the spinal cord, and recognized this as a distinct neurological disease, giving it the name "sclerose en plaques"[3]. The name is based on the appearance of disseminated demyelinated nervous tissue and axonal loss, involving destruction of the myelin producing cells (the oligodendrocytes) and loss of the myelin sheath covering the neuronal axons (crucial for action potential and nerve signal propagation), resulting in subsequent axonal damage [1].

1.1.1 Epidemiology of MS

There are about 2-2,5 million MS patients worldwide and MS has an asymmetrical distribution, with high prevalence in northern European and American populations[4]. There is a strong association between the prevalence of MS and global latitudes, the higher the latitude the higher the prevalence[5]. Some regions do not agree with the latitude gradient distribution, such as the indigenous Sami population in Scandinavia that has a low prevalence of MS. The prevalence of MS in Sweden is 188.9/100,000 individuals, with a female to male ratio of 2.35:1[6].

1.1.2 MS clinical manifestation and diagnosis

MS patients usually experience their first symptoms when they are 20 to 40 years old. After experiencing the first neurological episode the patient is considered as clinically isolated syndrome (CIS). However, when typical MS magnetic resonance image (MRI) lesions are detected with the absence of clinical signs and symptoms, the patient is considered to suffer from radiologically isolated syndrome (RIS). CIS and RIS patients will be followed over time up to see if the patients develop MS. MS has different courses; 85% of the patients have a relapsing-remitting (RRMS) course, during which the patient has alternating periods of neurological disabilities and recoveries or in other words periods of relapses and remissions. When the disabilities start to accumulate and there is no recovery, RRMS usually advances to secondary progressive MS (SPMS). In the primary progressive MS (PPMS) course, representing 10-15% of MS patients and usually with an older age of onset than the RRMS, the patient's condition worsens from the beginning and continues developing neurological signs and symptoms without periods of recovery[7] (Figure 1).

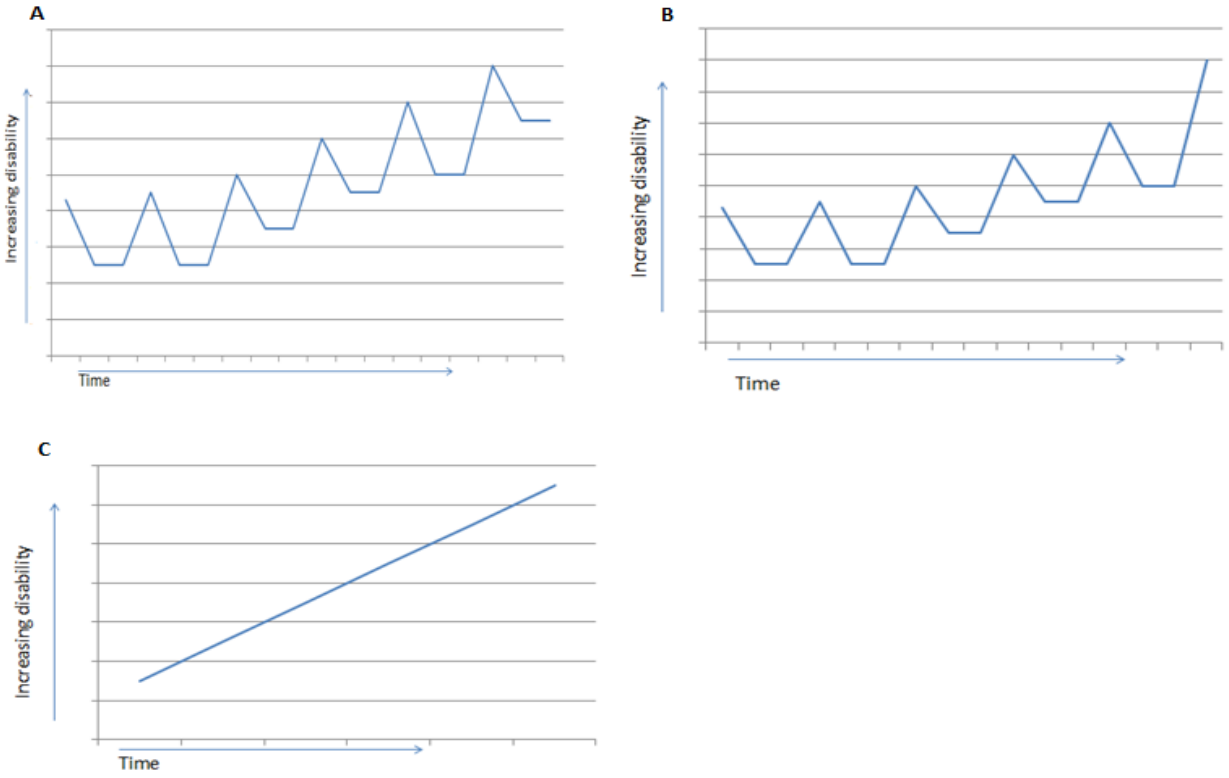

Figure 1. Different courses of MS; RRMS (A), SPMS (B), PPMS (C)

In order to make it easier for neurologists to decide on the right diagnosis earlier and at the same time avoiding a faulty diagnosis, the International Panel on the Diagnosis of Multiple Sclerosis formulated the McDonald criteria for MS diagnosis[8]. In summary, a diagnosis according to these criteria is based on the number of relapses, taking into account the dissemination in space and time, and aided by paraclinical methods, which are the detection of oligoclonal bands (OCBs) in the cerebrospinal fluid (CSF) with absence in the serum and magnetic resonance imaging (MRI) of the lesions. MS diagnosis could be reached without the help of paraclinical methods if the patient experiences two or more relapses which are separated in both time and space. Continuous reviews and revisions of the McDonalds criteria have been done since its presentation[9,10]. A grading system is used to assess the level of disability in MS patients, the Expanded Disability Status Scale (EDSS)[11]. The scale extends from normal (=0) to death from MS (=10).

1.1.3 MS central nervous system pathology

During the RRMS course the axonal demyelination is considered to result in blockade of the nerve signals and development of neurological symptoms. After cessation of the inflammatory process, limited remyelination of the axons results in remission[12]. Clinical symptoms usually mirror the site of the lesion in the CNS and the appearance of these lesions are commonly associated with the blood brain barrier (BBB) being breached by autoreactive immune cells, development of inflammation at different sites and demyelinated plaques[13]. These demyelinated plaques are the pathological hallmark of MS and they also contain limited preserved axons and astrocytic scars. They can randomly present within areas of the CNS, but certain areas are more inclined to be affected; optic nerves, brainstem, cerebellar white matter,

periventricular and subcortical white matter and the cervical segments of the spinal cord, and they often surround large or medium sized veins. Demyelination in MS has for a long time been thought to occur only in the white matter, considering MS a white matter disease, but the truth is that demyelination can appear extensively in the gray matter, especially in the cortex. In addition, neuronal and axonal loss and atrophy can be found in the demyelinated and non-demyelinated cortical gray matter areas even in a global manner. MS brains are also affected in a global manner, especially for SPMS and PPMS. Despite retaining myelin, the normal appearing white matter (NAWM) is chronically injured, characterized by diffuse inflammation consisting of small perivascular inflammatory cuffs and microglial activation forming microglia nodules, in addition to the presence of axonal injury independently of demyelination. These diffuse changes in the NAWM result in global brain atrophy [14]. Taking into consideration the focal demyelinated lesions in the white matter and also the gray matter and the diffused axonal loss and inflammation in the NAWM, one can conclude that MS is a CNS global phenomena involving neuro-inflammation and neurodegeneration.

1.1.4 Autoimmunity in MS

1.1.4.1 CD4 and CD8 T cells

Most experts in the MS immunity field agree with the hypothesis that MS pathogenesis involves activation of CD4 T cells in the periphery by self-myelin antigens, followed by activated cells crossing the BBB and initiating the inflammatory process. This hypothesis is supported both by the MS experimental model, experimental autoimmune encephalomyelitis (EAE), in which an MS like disease is induced by myelin specific CD4 T cells, and by genetic studies which show that *HLA* class II loci show the strongest association to MS risk. Even though there is a majority agreement on this hypothesis, there is still somewhat of a mystery as to where these CD4 T cells are activated and which antigens specifically activate them. These antigens could be self CNS antigens or cross-reactive antigens. Cross-reactivity occurs when foreign antigens, such as viral antigens, have similar peptide sequences to CNS antigens and might activate CD4 T cells in the periphery which subsequently may be exposed to CNS antigens resulting in reactivation of these cells. This is termed molecular mimicry. Several studies have shown the presence of these myelin autoreactive CD4 T cells in the circulation of MS patients and also in controls, but in much lower frequencies. These autoreactive cells were also shown to be able to secrete the inflammatory cytokine Interferon γ (IFN-γ). CD8 T cells are also implicated in causing CNS tissue damage, but they are not as well studied in the context of MS as CD4 T cells and are thought to be involved more in the chronic phase[15]. CD8 T cells are more abundant in the MS lesions than the CD4 T cells[16,17] and show clonal expansion[18]. Genetic studies also advocate the role of CD 8 T cells, as alleles in the *HLA* class I region have been associated with MS, *HLA-A*02:01* being protective and *HLA-A*03:01* increasing the risk [19,20]. This makes sense because neurons and axons express the major histocompatibility complex (MHC) class I molecules which are recognized by the C8 T cells.

CD4 T cells are classified to different functional classes or T helper lineages based on their cytokine profile. The most studied lineages are the T helper 1 (TH_1) and T helper 2 (TH_2). TH_1 produce IFN-γ, while TH_2 produce IL-4, IL-5 and IL-13[21]. Recently a new class of helper T cells has been identified that express IL-17, TH_{17} cells[22]. The autoreactive CD4 T cells involved in MS have for a very long time been believed to be only of the TH_1 type[23]. This is no longer the case, as lately the involvement of TH_{17} cells was also evident[24]. Furthermore, the number of TH_{17} cells in the CSF increases during relapse[25]. IL-17 producing cells were found to present in active and chronic MS lesions. These were CD4 T cells and interestingly also CD8 T cells, astrocytes and oligodendrocytes[26]. TH_{17} cells are able to start producing IFN-γ under certain conditions[27]. Therefore, one must take into consideration the plasticity of the T helper cells, them being able to switch from one class to another depending on the surrounding cytokine environment[28] , when characterizing the autoreactive T cells in MS.

1.1.4.2 γ/δ T cells

γ/δ T cells constitute approximately only 5% of the total T cell population in the blood circulation while in the skin and intestinal epithelium they are more common constituting 50% of the T cells. Moreover, unlike αβ T cells, they recognize protein and non-protein antigens and do not require antigen presentation by MHC molecules (no MHC restriction)[29]. γ/δ T cells are part of both the innate and adaptive immune system, hence classified as innate-like T cells[30]. During their development in the thymus they can commit to a fate of IL-17 or IFNγ secreting cells[31]. There is evidence of their involvement in MS pathology, earlier studies have identified expanded clones in acute MS plaques and CSF[30] and even recent studies have examined the involvement of γ/δ T cells in the inflammatory process of MS[32,33].

1.1.4.3 B cells

B cell involvement in MS autoimmunity has not been of great interest despite the presence of the Ig OCBs in the CSF and B cells forming lymphoid follicles in the meninges[34]. However, following the success of the anti-CD20 treatments, rituximab and ocrelizumab, more and more interest is directed towards B cell research in MS[35,36]. The role that B cells play in the immuno-pathogenesis in MS is currently proposed to be antibody independent, mainly due to the fact that the antibody producing plasma cells do not express CD20 and OCBs are not much effected by anti-CD20 therapies[37]. The suggested antibody independent mechanisms are antigen presentation and cytokine production. Antigen presentation to T cells is an important step for their activation and differentiation. B cells express both pro- and anti-inflammatory cytokines and depending on their balance they can activate or downregulate the immune response[38].

1.1.5 Treatments

Unfortunately, there is currently no cure for MS. However, the introduction of the immunomodulatory treatments, resulting in changes in the immune system to a more regulatory or suppressive state, has been successful in reducing the number of relapses and providing a better quality of life for RRMS patients.

1.1.5.1.1 Interferon beta

Interferon beta (IFNβ) is a polypeptide that is naturally present in the human body. It is a cytokine belonging to the type I interferons and is involved in the defense mechanism against viruses. Recombinant forms of IFNβ are the most widely used treatment for MS. They are IFNβ-1a; administrated intramuscularly or subcutaneously and IFNβ-1b; administrated subcutaneously. Treatment with IFNβ reduces the annualized relapse rate (ARR) by 30% and MRI disease activity[39]. For SPMS patients IFNβ treatment benefited only the reduction of relapse rate but not disability progression[40], while showing no effect for PPMS patients[41]. The mechanism behind its clinical effect is not quite understood, but it is believed that IFNβ shifts the cytokine milieu more to an anti-inflammatory state and reduces the crossing of the immune cells through the BBB[42].

1.1.5.1.2 Glatiramer acetate

Glatiramer acetate is a synthetic peptide that has some similarity with the sequence of myelin basic protein (MBP). It is administrated subcutaneously and its suggested mode of action is that when it binds to the MHC class II molecules, either regulatory T cells (Tregs) are activated which cross react with MBP activated T cells and suppress them, or by binding to the MHC class II, glatiramer acetate competes with the binding of myelin antigens and hence reduces the activation of effector T cells[43]. Clinical trials for glatiramer acetate have shown a 29% reduction in ARR and also reduction of enhancing MRI lesions for RRMS[44,45].

1.1.5.1.3 Teriflunomide

Teriflunomide is an orally administrated drug that exerts its immunomodulatory effect by inhibiting the mitochondrial enzyme dihydrooroate dehydrogenase which is essential for pyrimidine synthesis, thus inhibiting synthesis of pyrimidine and ultimately cell proliferation. How this benefits MS treatment is not totally understood, but it could be by reducing the proliferation of myelin autoreactive immune cells[46]. Its clinical efficacy is comparable with IFNβ [47].

1.1.5.1.4 Dimethyl fumarate

Dimethyl fumarate (DMF) is an orally administrated immunomodulatory drug that demonstrates a marked reduction in ARR and MRI disease activity[48]. Its mechanism of action is still not fully understood, but it is thought to act mainly through the anti-oxidative nuclear factor (erythroid-derived 2)-like 2 (Nrf2) transcriptional pathway[49]. Mechanisms independent

from the Nrf2 pathway have also been suggested[50]. Recently, it was reported that patients with a favorable response to DMF treatment have an increased reactive oxygen species production by monocytes[51].

1.1.5.1.5 Natalizumab

Natalizumab, is a humanized monoclonal antibody that binds to α4-integrin of the cell adhesion molecule VLA-4, blocking it from binding to its ligand VCAM1 on the surface of the endothelium. This prevents T cells in the peripheral circulation from crossing the BBB and gaining access to the CNS. Natalizumab is administrated intravenously and its clinical trials showed impressive results, reducing the ARR by 68%, the rate of disability progression by 54% and gadolinium enhanced MRI lesions by 92% compared to placebos[52]. Treatment with natalizumab is coupled with the risk of developing of a rare devastating viral CNS infection, progressive multifocal leukoencephalopathy (PML). PML occurs as a consequence of reactivation of the John Cunningham virus (JCV), when the immune surveillance of the CNS is compromised by natalizumab. The incidence of PML in patients who have had the treatment for two years is 1 per 1000 patients[53]. As precautions for PML development, the titers of anti-JCV antibodies are followed in patients treated with natalizumab for benefit–risk treatment decisions [54].

1.1.5.1.6 Fingolimod

Fingolimod is a structural analogue of sphingosine and it is the first oral drug approved for MS treatment. Sphingosine, when phosphorylated, binds to the sphinogsine-1 phosphate receptor type 1 (S1PR1) on the lymphocytes and induce their egress from secondary lymphoid organs. After phosphorylation, fingolimod binds to the S1PR1 which leads to internalization of the receptors, thereby preventing lymphocytes from exiting the secondary lymphoid organs[55]. Compared to placebo, fingolimod reduced the ARR by 48-55%, the rate of disability progression by 25-30% and gadolinium enhanced MRI lesions by 80% compared to placebo[56,57]. There are concerns for varicella zoster virus infection associated with fingolimod treatment[58]. Therefore, screening for previous infection and vaccination are recommended before treatment initiation. Fingolimod might cause transient bradycardia and atrioventricular block, thus when taking their first dose patients are monitored using electrocardiogram.

1.1.5.1.7 Rituximab

Rituximab is an anti-CD20 monoclonal mouse chimeric IgG1 that depletes CD20 B cells[59]. Currently it is approved for the treatment of B-cell lymphoma and rheumatoid arthritis, however the treatment of MS patients with rituximab is increasing in Sweden even though off label[60]. In Stockholm and Västerbotten counties, rituximab was found to perform better than natalizumab, fingolimod and DMF in regards to effectiveness and discontinuation (data not shown).

1.1.5.1.8 Other RRMS treatments

Inducing lymphocyte depletion and immune reconstitution is also an effective treatment strategy and this can be achieved by the drugs cladribine and alemtuzumab[61]. In addition, autologous haematopoietic stem cell transplantation has been adopted for treating patients with a high disease activity[62].

1.1.5.1.9 Progressive MS treatment

Options for treating the progressive course of MS is limited, however there are a number of ongoing and completed clinical trials of different strategies for preventing disability progression[63]. Two drugs which were promising for reducing disability progression are now approved for usage; ocrelizumab for PPMS[64] and siponimod for SPMS[65].

1.2 MS ETIOLOGY

MS is a complex disease in which both environmental and genetic factors are involved. The Rothman pie model[66] has been used to explain this complexity. The model visualizes the environmental and genetic factors, which are the contributing factors, as slices of a pie which together make up sufficient cause for the disease to develop ("the whole pie"). Different sufficient causes can lead to development of the same disease and the same contributing cause could be shared in different sufficient causes (Figure 2).

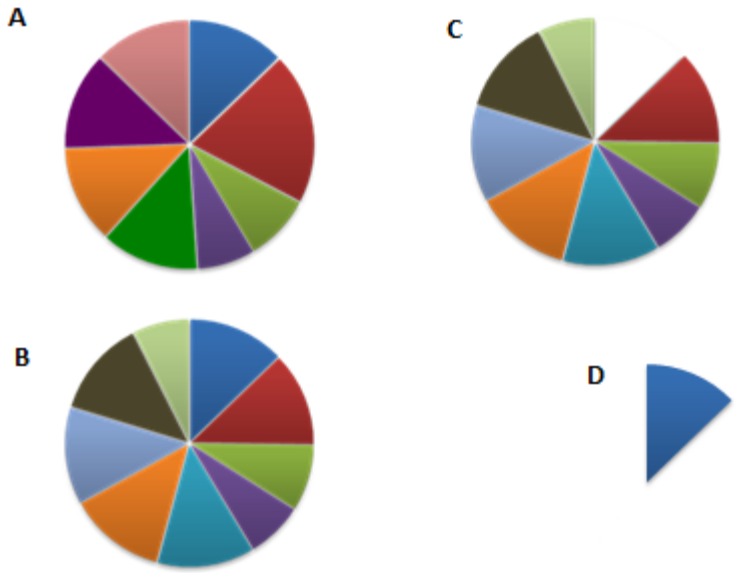

Figure 2. The pies A and B are sufficient causes, while C is not sufficient because it misses the slice D (contributory cause).

1.2.1 Environmental risk factors

Several environmental factors have been associated with the development of MS[67]. High vitamin D levels were found to be protective against MS and also reduce the hazard of relapses and MS brain lesions[68,69]. Cigarette smoking, as is the case with many other diseases, is a risk factor for MS[70]. In addition, MS patients who continue smoking after diagnosis have a worse prognosis than those who quit[71]. Working night shifts before the age of 20 was also identified as an MS risk factor[72]. Infectious mononucleosis, an infection caused by Epstein Barr virus (EBV) in adolescents and young adults, is also associated with the risk of developing MS[73].

1.2.2 Genetic risk factors

Genetic inheritance of MS susceptibility is of important for the development of MS. This has been proven by familial risk and twin studies of MS patients and their relatives which reported high concordance rates for twins and increased risk for relatives of MS patients[74,75]. A meta-analysis of the published familial recurrence risk data found that the age adjusted risk (AAR) was 18.4% for monozygotic (MZ) twins 4.61% for dizygotic (DZ) twins, 2.68% for siblings and 1.45% for parents[76]. A recent study on the Swedish population has shown slightly lower AAR, with 17.26% for MZ twins, 1,92% for DZ twins and 2.55% for siblings[77].

In 1973, the first genetic locus associated with MS susceptibility was identified by a Danish group in the HLA class II region, then called the *HLA7* [78]. With the development of genotyping techniques this association was further confirmed and the haplotype *DRB1*1501-DQA1*0102-DQB1*0602* was found to be associated with MS susceptibility[79]. Candidate gene association studies have identified a number of MS associated genes, including *IL2Rα* and *IL7Rα*[80,81].

1.2.2.1 MS Genome wide associations studies

Based on the common disease common variant hypothesis, genome wide associations studies (GWAS), mainly in a case-control design, have been performed to identify susceptible common variants to MS. The advancement of microarray technology that allowed genotyping of SNPs covering the whole genome and the provision of thousands of patient samples through international collaborations (the international MS genetic consortium; IMSGC) have contributed to the success of GWAS in MS in the past decade[82], discovering so far more than 200 associated variants outside the MHC region with relatively small odds ratios reflecting a small effect size of these variants[82]. Furthermore, these large numbers of associated variants explain only about 30% of the heritability of MS[82]. Hence the term "missing heritability" has been introduced also in to the field of MS genetics[83], as well as in other complex autoimmune diseases[84]. There have been several suggestions of how to further the understanding of the genetics behind complex disorders and to facilitate the interpretation of the involvement of so many genetic variants with minor effects on disease development. These include the analysis of genetic pathways, study of rare genetic variants rather than common variants, more sophisticated analysis methods and whole genome or exome sequencing initiatives[85].

1.2.2.2 Different types of genetic variants

Variations in the human genome could range from a single nucleotide to a whole chromosome variation. Single nucleotide variants can be common or rare with the common variants have a frequency of 1% or more in the population. Common single nucleotide variants are known as single nucleotide polymorphisms (SNPs) occurring as mostly two different alleles of different frequencies. SNPs are widely studied as associated with the risk of developing a specific disease. Variations that are more than one nucleotide up to 1kb in the form of deletions or insertions are called indels. Furthermore, duplication or deletions of a DNA segment ranging from a couple of thousands to a few millions base pairs are described as copy number variations (CNVs).

When the variation in the coding DNA sequence is different from the reference genome it is considered a mutation. Mutations can be hereditary or acquired. Hereditary mutations are inherited from the parents and is present in all cells, while acquired mutation is usually a consequence of an environmental effect or failure of DNA damage repair and is present in certain cells[86].

1.2.2.3 Somatic mosaicism

Differences in the DNA sequence between individuals is always expected but not within the same individual between different tissues. The presence of such genetically distinct populations of somatic cells within an individual is called somatic mosaicism. This can be on the level of single nucleotides or whole chromosomes. An obvious example of somatic mosaicism is neoplasia, but there are also Mendelian disorders which may be characterized by somatic mosaicism, such as Fanconi anemia and Hemophilia A[87]. Although perhaps provoking, somatic variations in the form of CNV have also been reported in healthy individuals[88,89]. A form of somatic variation which is well recognized is the rearrangements in T cell receptor (TCR)s and B cell receptor (BCR)s that provides the diversity in T and B cells, respectively, which contributes to the diversity of the adaptive immune system. It was actually more than 60 years ago in 1957 that Burnet described this mechanism when he hypothesized that lymphocytes undergo 'randomization' of the genes coding the immunoglobulins in his clonal selection theory for the formation of antibodies[90].

There have been previous attempts to identify genetic variations specific to the brain in neurological diseases. CNV in the brain, but not in the peripheral blood, have been described in schizophrenia patients, but also in controls[91]. Another study by Pamphlett et al[92] identified CNVs which were unique to the brain of sporadic ALS patients when compared to their blood.

1.3 TCR

TCRs are the transmembrane molecules through which the T cells recognize the antigens. They are a heterodimer of either an α and β or γ and δ chains that are produced during the maturation of T cells in the thymus through the rearrangement of multiple polymorphic genes to eventually produce either an αβ or γδ TCR[93]. The α and γ chains are coded by variable (V), joining (J) and constant (C) genes and the whole region covering these genes are named the TCR α (TRA) and TCR γ (TRG) loci on chromosomes 14 (14q11) and 7 (7p14), respectively. The β and δ chains are coded by V, diversity (D), J and C genes. The TCR β (TRB) locus is situated on chromosome 7 (7q34) while the TCR δ (TRD) locus is situated within the *TRA* locus on chromosome 14 (14q11). Figure 3 illustrates reassembly of V(D)J genes in the process of TCR rearrangement. The diversity of the TCRs is mainly determined by the selection and rearrangement of the V(D)J genes, as the C genes are highly homologous. Furthermore, the region where the V(D)J genes are assembled together undergo additional nucleotide deletions or insertions producing a more unique sequence called the complementarity determining region (CDR)3. Hence, each T cell with a unique CDR3 sequence represents a distinctive T cell clone i.e. if two TCRs are coded by the same V(D)J genes they could still belong to two different clones. For every clone, each α or γ chain combine with a single β or δ chain, respectively. Therefore, determining the sequence of one of the chains is sufficient to identify the clone[94]. In project II of this thesis we have studied the T cell clonality through the analysis of the TRB repertoire by sequencing the TRB CDR3 sequence[95].

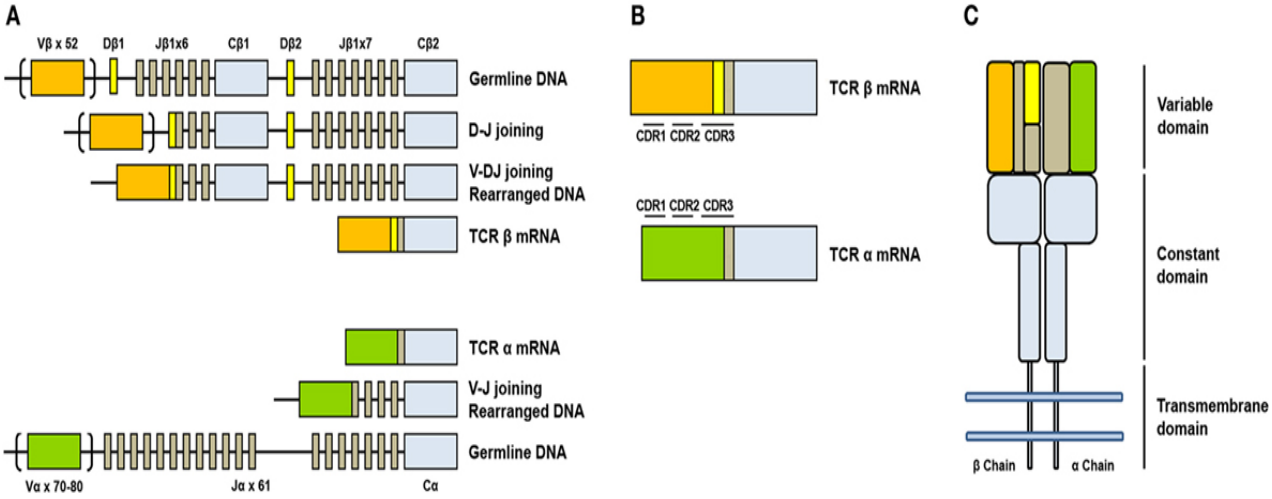

Figure 3. Illustration of the αβ TCR rearrangement. A) Somatic reassembly of TRB and TRA loci. B) Coding TRB and TRA mRNA. C) αβ heterodimer. [Figure reproduced from "De Simone, M., Rossetti, G. & Pagani, M. Single Cell T Cell Receptor Sequencing: Techniques and Future Challenges. Frontiers in Immunology ***9****, doi:10.3389/fimmu.2018.01638 (2018)"][96]*

1.4 TCR DATABASE

There are several databases for TCR sequences available for the public, e.g. Mc-PAS-TCR[97] and VDJdb[98]. TCR databases are a great resource of information, as they provide details of the reported CDR3 sequences and their target antigens if available and in which pathology or organ they were detected. In the context of MS, TCR databases are of great help to the scientific community, as the target antigen of the autoreactive T cells is still a mystery and the access to the target organ, the CNS, is invasive and excluding CSF samples, usually postmortem[99]. TCR databases may hence limit the number of CNS samples required in individual studies and facilitate the identification of overlap between different studies. Furthermore, databases may improve understanding of the role of molecular mimicry in MS in case of the same CDR3 sequence being reported to have different target antigens.

1.5 BIOMARKERS

The term biomarker is a commonly used term in the medical field, but what is a biomarker? One definition of a biomarker is "a characteristic that is objectively measured and evaluated as an indicator of normal biological processes, pathogenic processes, or pharmacologic responses to a therapeutic intervention."[100]. Biomarkers are useful for disease screening, diagnosis, disease prognosis and activity monitoring, assessing treatment effect and response and treatment tailoring.

One of the oldest biomarkers in MS is the OCBs which are oligoclonal IgG bands present in the CSF while absent in the plasma. The presence of OCBs is useful for the differential diagnosis of MS[101] and even though OCBs are present in 95% of MS patients there has been hesitation in accepting it as a diagnostic tool[9] until recently. After studies emphasized the significant role of OCBs in the diagnosis of MS[102], it was recommended by the 2017 revisions of McDonald criteria, as a criterion to fulfill the dissemination in time in CIS patients with an established dissemination in space[10]. MRI is important for the diagnosis and monitoring of MS. MRI follow ups give information on new white matter lesions, the disease activity by measuring active inflammation and global CNS changes[103]. Neurofilament light chain (NfL), a component of the cytoskeleton of neurons, is a potential biomarker in MS. CSF levels of NfL is associated with the activity[104] and progression of MS[105]. The serum and CSF levels of NfL are (luckily) correlated providing a more convenient access[106]. Treatment of MS patients with natalizumab reduced the levels of NfL in the CSF[107], however it increased serum NfL before PML development, therefore NfL was suggested as a potential marker for early PML development[108], while, finoglimod reduced the NfL levels in the CSF[109] and plasma[110]. Therefore, NfL is a promising biomarker for MS progression and treatment response. Anti-drug antibodies (ADAs) are a product of the immune system against biopharmaceuticals entering the body. A small portion of the patients treated with biopharmaceuticals will develop neutralizing ADAs (nADA) reducing the efficacy of the treatment[111]. Monitoring of ADAs in

connection with MS treatment is now common clinical practice in some countries. Furthermore, with the current line of management of individualized tailored medicine, the need for biomarkers in MS is rising. We need biomarkers which we can access without causing a great discomfort to the patient and being cost-effective at the same time.

1.6 INTERLEUKIN-2, 6 AND 7 RECEPTORS

The discovery of genetic associations is a foundation to identify the affected pathways and mechanisms in MS. Functional studies have shown that some of the associated variants can influence the expression of the genes. This is the case of rs6897932-C and rs1800693-G which promote alternative splicing of *IL7R* and *TNFRSF1A* transcripts, respectively, and the formation of the soluble form of their proteins[112,113]. In addition, in vivo studies have shown that the associated variants in *IL7R* and *IL2R* effect their soluble protein levels[112,114].

In two of the studies in this thesis, we have looked at the corresponding proteins of genes which have shown association with MS susceptibility and examined their biomarker potential.

1.6.1 IL-2Rα

Four decades ago, interleukin 2 (IL-2) was discovered[115] as a crucial factor for T lymphocyte proliferation. Interleukin 2 receptor alpha (IL-2Rα; CD25) binds to IL-2 with a low affinity and does not transfer IL-2 signals without being in the trimeric complex of IL-2R with IL-2RB (CD122) and the common gamma chain (IL-2rγ or γc)[116]. Presence of a soluble form of IL-2Rα (sIL-2Rα) in vitro was first described in 1985 by Rubin and colleagues[117]. Biochemically they described sIL-2Rα as a glycoprotein 10 Kd shorter than the membrane bound receptor and readily binding IL-2[118], hinting to an inhibitory effect of the soluble receptor. They also concluded that sIL-2Rα is mainly produced by proteolytic cleavage of the membrane bound receptor[119]. However, the production of the IL-2Rα without the transmembrane domain through an alternative splice variant of the *IL2RA* mRNA has been also demonstrated[120]. Therefore cells expressing the membrane receptor would be expected also to secrete the soluble receptor including T cells, B cells and monocytes[121]. Many signaling pathways of IL-2 are linked to the beta subunit, including Janus kinases (JAK) and the signal transducer and activator of transcription (STAT)[122] (Figure 4).

After the identification of sIL-2Rα, many research labs studied its levels in different inflammatory contexts, including infections, post transplantation, malignancies and autoimmune diseases. Increased levels of serum sIL-2Rα were detected in patients with rheumatoid arthritis[123], type 1 diabetes, Crohn's disease, ulcerative colitis, systemic lupus erythematosus and MS[121]. Hence, detecting increased levels of sIL-2Rα could be a surrogate for immune activation. In healthy individuals serum and urine soluble IL-2Rα levels seem to be higher in children than in adults[124,125], but whilst higher in the plasma of elderly than in young adults[126] there seems to be no difference between females and males.

Both candidate and genome wide associations studies (GWAS) have implicated SNPs within the *IL2RA* gene to be associated with MS susceptibility[80, 127, 128]. Using immunochip to study the immune related loci in a GWAS, the A allele of the SNP rs2104286 in the intron between exon 1 and 2 of the *IL2RA* gene was the non-*HLA* allele with the highest associated risk for MS[128]. Cerosaletti et al showed that having the MS risk haplotype for rs2014286 is correlated with reduced IL-2 signaling in CD4 T cells expressing high levels of IL-2Rα and that this IL-2 reduced signaling is correlated with sIL-2Rα levels[129].

1.6.2 IL-6R

Interleukin 6 (IL-6) is a multifunctional cytokine that has both inflammatory and anti-inflammatory functions[130] and is a major inducer of acute phase protein production by hepatocytes[131]. To exert its function, IL-6 binds first to the IL-6 receptor (IL-6R) on the cell membrane and then this complex binds to the signal transducing molecule glycoprotein 130 (gp130) which then dimerizes and starts the signaling[132] . This is known as the classic signaling of IL-6. IL-6 intra-cellular signaling passes through the JAK/STAT pathway, mainly STAT3 and secondarily STAT1 as well as the mitogen activated protein kinase (MAPK) signaling pathway[133]. The classic signaling of IL-6 is limited to certain types of cells which express the IL-6R, including hepatocytes, megakaryocytes, monocytes, macrophages, B cells and subtypes of T cells[134]. The signaling transducing molecule gp130 is ubiquitously expressed and cells that do not express IL-6R receive IL-6 signaling through first forming a complex with a soluble form of IL-6R (sIL-6R) and then binding to gp130 on the surface of the cell[135]. Rose-John and Heinrich named this alternative signaling "IL-6 trans-signaling"[136]. sIL-6R is produced both by mRNA alternative splicing[137] and protease cleavage by metalloprotease ADAM17[138] (Figure 4). The classic signaling of IL-6 through the membrane bound IL-6R transfers the anti-inflammatory effects of IL-6, while the IL-6 trans-signaling transmits the pro-inflammatory effects[130]. Kishimotos group managed to show also the existence of a soluble form of gp130 (sgp130) and that this sgp130 inhibits IL-6 trans-signaling by binding to the IL-6/sIL-6R complex and preventing it from binding to the membrane bound gp130[139] (Figure 4). sgp130 is mainly produced by alternative splicing of the mRNA. IL-6 levels in the human peripheral circulation are about 1-5 pg/ml, which is 1000 times lower than the levels of sIL-6R at about 50 ng/ml, while sgp130 levels are about 400 ng/ml[140]. IL-6 trans-signaling blocks the development of the autoimmune inhibitory Tregs cells from naïve CD4 T cells[141] and favors the development of autoimmune inducing TH_{17}-cells instead[142, 143].

In rheumatoid arthritis, the synovial fluid from arthritic patients showed increased levels of sIL-6R which corresponded to the advanced stages of the disease and leukocyte infiltration[144, 145] . The involvement of IL-6 in the pathogenesis of MS has also been of interest, from showing elevated expression in MS lesions, in CSF from MS patients and in peripheral blood mononuclear cells[146,147], to failure of the development of MOG induced EAE in IL-6 deficient mice[148]. Conflicting results of higher and normal levels of IL-6 in the peripheral circulation of MS patients have been reported[149,150]. Regarding the studies of the receptors of IL-6 in MS, Padberg et al showed higher levels sIL-6R and sgp130 in the serum of MS patients compared

to healthy controls. IL-6 involvement in MS autoimmunity could be through both T cell and B cells. IL-6 is important for B cell differentiation and IgG production; hence its old name "B cell differentiation factor"[151] and B cells from MS patients were shown to express high levels of IL-6 which could be lowered by the B cell depleting drug, Rituximab[152]. Furthermore, IL-6 secreted by B cells in the context of toll like receptors (TLR) activation increases the proliferation of effector T cells in vitro[153]. Effector T cells from RRMS patients have been shown to develop resistance to Tregs through the IL-6 signaling pathway[154]. Inhibition of IL-6 signaling in the EAE model using IL-6r antibodies reduced the onset of the MOG induced EAE and the development of MOG specific TH_{17} and TH_1 cells[155]. Based on these findings, tocilizumab, a humanized anti-IL-6R monoclonal antibody, has been suggested for MS treatment.

1.6.3 IL-7Rα

Interleukin 7 (IL-7) is an essential factor to T lymphocytes from their maturation and differentiation within the thymus[156] to the survival and homeostasis of naïve and memory T cells in secondary lymphoid organs and peripheral circulation[157]. IL-7 production is not by the T lymphocytes themselves (as in the case of IL-2 and IL-6) but is produced mainly by stromal cells at primary and secondary lymphoid organs[158]. The signaling of IL-7 goes through binding to a heterodimeric receptor composed of Interleukin 7 receptor alpha (IL-7Rα; CD127) and the common gamma chain. This initiates mainly the JAK/STAT (STAT5) and phosphoinositide 3-kinase (PI3K) signaling pathways[159] (Figure 4). In 1990, a soluble form of the IL-7Rα (sIL-7Rα) that can bind to IL-7 was identified[160]. It is mainly produced by alternative splicing of the *IL7RA* transcript[161], removing exon 6 which codes for the transmembrane domain. The concentration of sIL-7Rα in the peripheral circulation is about 1000 fold higher than its ligand, IL-7, ranging from 0,3-8,4 pg/ml[162,163]. The function of this soluble receptor still needs to be better understood. One study proposed an inhibitory effect to IL-7 activity[164] and recently another study showed that sIL-7Rα increases the availability of IL-7 and enhances its activity [112].

Regarding the involvement of IL-7 in the process of autoimmunity, evidence point to its implication in type 1 diabetes and MS[165]. IL-7 signaling has also been shown to enhance the response of autoreactive T cells against myelin antigens from MS patients[166]. In addition, the expression of both IL-7 and IL-7Rα proteins was found to be higher in the CSF of MS patients than in controls with other non-inflammatory neurological diseases[81]. Interestingly, actively suppressive Tregs are low expressers of IL-7Rα[167]. In addition, the major allele (C) for the SNP rs6897932, located in exon 6 of *IL7RA*, was the first allele outside the HLA region found to be associated with the risk for MS[81]. Later on, large cohort MS GWAS have confirmed the association of this SNP with MS risk[127]. Furthermore, presence of this MS associated allele is associated with increase in expression of the alternative transcript of the *IL7RA*, missing the transmembrane domain[168]. Likewise, an increase in the sIL-7Rα was associated with the

rs6897932 C allele in an allele dose dependent manner in MS patients, which was also accompanied by an increase in IL-7[112].

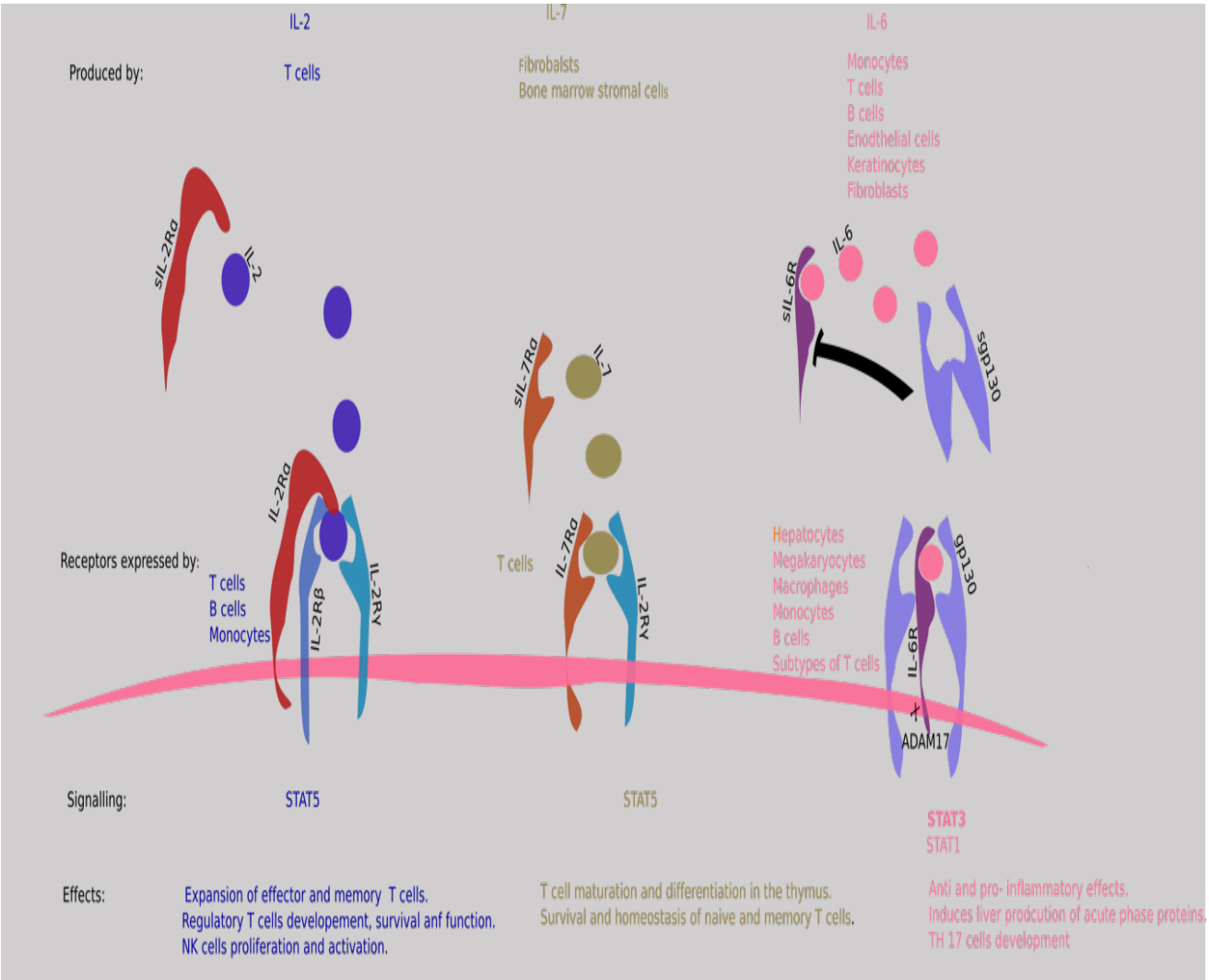

Figure 4. Schematic view of IL-2 ,7 & 6 and their corresponding receptors, intracellular signaling molecules and effects.

1.7 PEBP1

This protein has multiple names, each one describing its different function, and is expressed in multiple tissues[169]. Phosphatidylethanolamine binding protein 1 (PEBP1) binds to the phospholipid phosphatidylethanolamine located at the cytoplasmic side of the plasma membrane. It is also known as hippocampal cholinergic neurostimulating peptide precursor protein (HCNP-pp), acting as a precursor of HCNP which has been shown to enhance acetylcholine synthesis in vitro[170] . HCNP-pp is expressed both in the hippocampus and amygdala. Its expression was described to be lower in the hippocampus in Alzheimer patients compared to non-demented patients[171], while elevated in the amygdala of females with major depression[172]. PEBP1 is also involved in the intracellular signaling, it inhibits the mitogen activated protein kinase (MAPK) pathway by binding and inhibiting RAF1 protein kinase,

hence the name RAF1 kinase inhibitor (RKIP)[173]. This inhibition of the MAPK pathways is essential for tumor suppression[174]. PEBP1 is also expressed in T cells and long term exposure to its derivative HCNP suppresses choline acetyltransferase expression and decreases the amount of acetylcholine[175], conflicting its enhancement of acetylcholine synthesis in the hippocampus[170]. Hence, HCNP may act as a regulator in the lymphocytic cholinergic system that plays a role in the differentiation and proliferation of T cells[176].

2 THESIS AIMS

The general aim of this thesis is to explore genetic variants associated with MS and to further employ this knowledge to study potential MS biomarkers.

I. To identify genetic variants associated independently to the progressive course of PPMS and to the relapsing course RRMS using exome sequencing data.

II. To identify somatic mosaic sub-set of cells using CNVs that predispose specific immune cell to cross the BBB.

III. To explore the potential role of the gene products of the MS associated variants *IL7RA* and *IL2RA* and *IL6ST* as biomarkers when monitoring MS treatment.

IV. To utilize protein microarrays to identify potential biomarkers for MS treatment using serial plasma samples from MS patients.

3 METHODOLOGICAL CONSIDERATIONS

3.1 PATIENTS AND CONTROLS SAMPLES

All the samples collected from patients and controls were part of cohort studies approved by the Swedish ethical review authority.

3.1.1 GEMS

Genes and environment in multiple sclerosis (GEMS), is a population-based case-control study where prevalent cases are identified from the Swedish MS registry. The aim of GEMS is to study the interactions between genes and environmental factors in MS[177]. Controls are selected to match the patient's age, sex and residential area.

3.1.2 EIMS

Epidemiological investigation of multiple sclerosis is a population-based incidence case-control study to identify incident cases of MS[70].

3.1.3 STOPMS I & II

Stockholm prospective assessment of MS (STOPMS) is an ongoing prospective study of the long-term development of MS in newly diagnosed MS patients and patients with neurological symptoms at the Karolinska University Hospital[178]. Controls are also recruited within the same hospital.

3.1.4 IMSE I, II & V

The Immunomodulation and MS Epidemiology (IMSE) I, II and V, are post marketing surveillance studies of Tysabri (natalizumab)[179], Gilenya (fingolimod)[180] and Tecfidera (Dimethyl fumarate) treatments in Sweden, respectively. The main aim of the IMSE studies is to evaluate the treatments safety and efficacy. In addition they facilitate studying the association between genetic variants and blood markers with disease activity, disability outcomes and side effects during treatment.

3.2 CLINICAL DATA

Patients' clinical data were obtained from the Swedish MS registry[181] and included information regarding their disease activity; relapses and brain and spinal cord magnetic resonance imaging (MRI) findings, and disease worsening and progression; expanded disability status scale (EDSS), multiple sclerosis severity score (MSSS), multiple sclerosis impact scale (MSIS29) and the symbol digit modalities test (SDMT).

3.3 PAPER I

3.3.1 SNPs and indels calling

The GATK best practices workflow for germline short variants discovery v3.6 was applied for SNPs and indels calling[182]. The workflow comprises of three major steps, first a preprocessing step of aligning or mapping the raw sequencing reads to the reference genome, here we used the hg19 reference genome, followed by marking of the duplicate reads, then finally recalibration or correction of the quality or confidence scores for each base provided by the sequencing machine producing a BAM file per sample. From this BAM file the second step of variants calling, SNPs and indels calling, was proceeded using the HaplotypeCaller in GVCF mode to produce the intermediate file, GVCF file. Then multiple GVCF files from multiple samples were consolidated or combined creating a directory containing a GenomicDB datastore. Creating this directory is important for speeding up the following joint genotyping step using the joint genotyping tool, GenotypeGVCFs, outputting a combined genotyped multi-sample VCF file. Then finally a filtering step, the Variant Quality Score Recalibration (VQSR) step, is done by assigning a quality score called the variant quality score log-odds (VQSLOD) to each variant calculated from a Gaussian mixture model based on highly validated datasets. Using the VQSLOD specified threshold variants are divided into quality tranches that can be used to filter the variants. Ultimately a final VCF file ready for downstream analysis is produced.

3.3.2 CNVs calling from exome sequencing data

In addition to calling SNPs and indels from the exome sequencing data we called CNVs using the CLAMMS tool[183]. The tool first divides exome capture regions into equally sized windows or regions and filter out regions with extreme GC content. Then the coverage values for each sample are normalized individually. CLAMMS, using the coverage from a reference panel of samples, performs a mixture model fitting each window to model its expected coverage distribution. Then finally CLAMMS applies a hidden Markov model using the normalized coverage values for the individual samples and the distributions from the fitted model to call the CNVs.

3.3.3 Functional and clinical variants annotation

Functional annotation of the called variants was performed using the ANNOVAR tool[184]. It gives information such as if the variant could change the amino acid sequence and if that might result in loss of function of the protein. For clinical relevant information on the variants, such as if the variant has been previously reported to be associated to a disease and with a deleterious effect or not, we have mined the ClinVardatabase[185].

3.4 PAPER II

3.4.1 Array CNV analysis

For whole genome CNV analysis we used a DNA microarray which is designed specifically for CNV analysis, the CystoScan HD array (Affymetrix). This microarray includes approximately 2.7 million markers, where 750,000 are SNPs and 1.9 million are non-polymorphic probes, with intragenic and intergenic marker spacing of 880 and 1,737 base pairs, respectively[186]. Normalization of the raw probe intensities to a reference panel and the paired peripheral blood (PB) and CSF CNV analysis were performed using the Nexus Copy Number software (BioDiscovery Inc, Hawthorne, CA). A threshold of a minimum five consecutive probes was used to call a CNV.

3.4.2 Taqman copy number analysis

To validate the CNVs identified from the genome wide screening we used Taqman copy number qPCR assays. For each identified CNV a set of probes, were selected to target its center and upstream and downstream genomic regions. To each reaction mix, a reference assay targeting the telomerase reverse transcriptase (*TERT*) gene was also included. The *TERT* gene is known to have two copies in a diploid genome, hence it can be used to normalize the target assay. CT values were imported to the CopyCaller™ Software (Applied Biosystems) to calculate the copy numbers of the target genes in the paired samples from each individual, once specifying PBMC sample as a calibrator.

3.4.3 TCR sequencing

We used high throughput TRB sequencing to investigate the TCR repertoire in paired CSF, $CD4^+$ and $CD8^+$ T cells. The LymphoTrack® TRB assay- MiSeq® kit (72250009, Invivoscribe) was used for library preparation, including multiplex primers that target the conserved V_β and J_β regions. Using the MiSeq Reagent Kit v2 (MS-102-2003, Illumina), paired-end 2x250 sequencing was ran on the Illumina MiSeq platform. Each run composed of eight samples per flow cell, including three paired samples from two patients plus positive and negative controls. The raw FASTAQ data was then imported to MiXCR software[187] to align the reads to the reference gene and then assemble the clonotypes identifying the CDR3 sequences of each clone. Downstream repertoire analysis was done using VDJtools[188].

3.5 PAPER III

3.5.1 ELISAs for the quantification of sIL-7Rα, sIL-2Rα, sIL-6R and sgp130

To measure the levels of sIL-7Rα in plasma we developed an in-house sandwich ELISA. In short, a monoclonal anti-human IL-7Rα (R&D systems) was used to coat a 96 well plate and after an overnight room temperature incubation period the plate was washed and blocked. The next day, after washing and blocking the plate, 1:20 diluted plasma samples and 7-point standard curve, prepared using recombinant human IL-7R*α* Fc chimera (R&D systems), were added to the plate. The bound IL-7Rα was detected using a biotinylated polyclonal anti-human

IL-7Rα (R&D systems). After adding streptavidin-HRP and its substrate a color developed. The developed color is in proportion to the amount of IL-7Rα and its intensity was measured on a spectrophotometer using 450nm filter. To measure sIL-2Rα, sIL-6R and sgp130, commercially available ELISA kits (R&D systems) were used. All samples from the same patient were included in the same ELISA plate to avoid inter-plate variations.

3.5.2 Genotyping data

MS associated variants were genotyped using the MS replication chip, a customized Illumina array developed for the IMSGC[189].

3.6 PAPER IV

3.6.1 High throughput proteins measurements

In paper IV, we utilized the high throughput multiplex affinity array, antibody suspension bead array, to measure the levels of 59 proteins in serial plasma samples targeted with 90 antibodies. These antibodies had been generated as part of the HPA project[190] and we adopted a previous protocol by Drobin et al with slight modifications [191]. Briefly, antibodies were coupled to color-coded magnetic beads and then combined in a suspension buffer creating the bead array. Plasma samples were labelled with biotin after dilution 1:10 in phosphate buffer saline. After labelling the samples were diluted 1:16 in an assay buffer and heat treated for 30 minutes at 56ºC and then added to the bead mixture distributed into 384-well plates. Then the plate was washed and streptavidin conjugated fluorophore added. Lastly, the plate was measured in a FlexMap3D instrument (Luminex corp.) and median fluorescence intensity (MFI) was reported for each bead identity.

3.6.2 ELISAs for antibodies specificity validation

The specificity of the PEBP1 antibody, HPA008819 (Atlas antibodies), was validated using an indirect and a sandwich ELISA. The RTN3 antibody, HPA015649 (Atlas antibodies), specificity was validated using an indirect and an inhibition ELISA. For more details regarding the ELISAs please refer to the materials and methods section of paper IV.

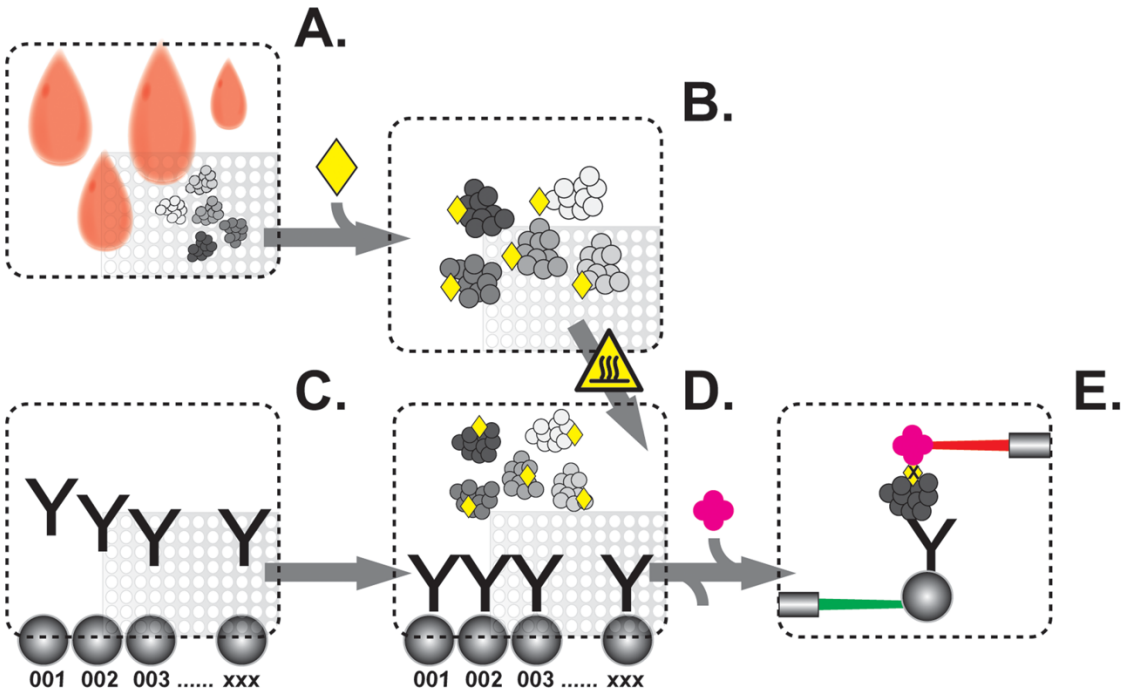

*Figure 5. Illustraion of the protocol of the antibody suspension bead array. A) Plasma samples are distributed into a microtiter 384-well plate. B) Biotin labelling of the proteins in the sample. C) Antibodies are coupled to color-coded magnetic beads and then mixed together to form the suspension bead array. D) Samples are heat treated and then combined with the beads mixture. E). The array is measured in a Luminex FlexMap3D instrument using two lasers to measure the intensity values (red laser) and to identify the bead with the coupled antibody (green laser). [Figure reproduced from "Darmanis, S. et al. Identification of candidate serum proteins for classifying well-differentiated small intestinal neuroendocrine tumors. PLoS One **8**, e81712, doi:10.1371/journal.pone.0081712 (2013)."]*[192]

3.7 STATISTICAL METHODS

3.7.1 Association tests for exome sequencing data

Association tests for the common variants obtained from the exome sequencing data were performed using the logistic Wald test on EPACTS version 3.3.0[193]. To correct for sex and relatedness, 20 principal components were obtained from the kinship matrix generated by the vcf2kinship tool that is part of the Rvtest package[194]and added to the logistic regression model. A p-value threshold of 5.5×10^{-7} and 5.5×10^{-5} was selected for an exome-wide significant and suggestive associations, respectively.

Association tests for rare variants are usually performed on a group of variants within a gene instead of single variants separately so as to increase the statistical power, known as gene-based association tests[195]. There are several methods for gene-based rare variant association tests and in paper I we have applied two methods. One is a burden test called the combined multivariate and collapsing (CMC) method[196], which assumes that all variants within the gene have the

same effect and their information is collapsed into a single score that is used for the association test. The second test we used is a combination of two methods, the burden test and the variance-component test (SKAT test) that takes into account that variants in the gene could have different effects on the trait (increase or decrease the risk), called the SKAT-O[197]. Genes with a p-value $<0.1x10^{-5}$ were considered to be significantly associated.

3.7.2 Frequency distribution testing

In paper II to test for the difference in the frequency distribution of expanded T cell clones between patients during relapse or remission we used the Pearson's Chi-squared test.

3.7.3 Linear mixed effect model

To test for the changes in the protein levels in a longitudinal manner during a period of approximately 24 months of treatment we used the linear mixed effect model. The benefit of using this model is that it deals with the random effect of having multiple measurement from the same individual while taking into consideration the fixed effects in the individual e.g. age and gender. We used this model in papers III and IV using the R package "LmerTest"[198].

3.7.4 Correcting for multiple testing

In paper IV, we utilized a multiplex protein array profiling 59 proteins, which necessitates the correction for multiple testing. Therefore, Bonferroni correction was applied to adjust the p values using the "multitest" package[199].

4 RESULTS AND DISCUSSION

4.1 PAPER I

We have utilized whole exome sequencing data from RRMS, PPMS and population-based control (PBC) to identify genetic variants that are associated to each course of MS independently, i.e. RRMS and PPMS. The association tests were conducted on three levels, PPMS vs PBC, RRMS vs PBC and PPMS vs RRMS.

The association of HLA class II genes (*HLA-DRB1*, *HLA-DQA1 and HLA-DQB1*) was replicated in both PPMS and RRMS. All significant associations in the HLA region disappeared after correction for HLA-DRB1*15 and A*02. Noticeably, the known protective association of the class I allele HLA-A*02 for RRMS[19] was not detected for PPMS. Outside the HLA region, when comparing PPMS to PBC, the locus of the pseudogene (*RP11-693J15.3*) on chromosome 12 carried one significantly and three suggestively associated indels to PPMS risk. Two nonsynonymous SNPs in the *VCAN* gene on chromosome 5 were also suggestively associated to PPMS. *VCAN* codes for the protein versican that belongs to the family of proteoglycans are one of the major constituents of brain extracellular matrix. In the edges of active brain white matter plaques of MS patients, versican was found to be increased [200].

When comparing RRMS to PBC outside the HLA region, we identified novel suggestive associations in the *CCL25, JMJD1C*, *LYZL2* and *TIMM44* genes to RRMS. A replication of these novel associations is desirable. CCL25 is a chemokine that is involved in T cell maturation and their recruitment to the intestine[201]. The expression of CCL25 was observed to be increased in the synovial fluids of rheumatoid arthritis patients and its antagonists are studied for the treatment of irritable bowel disease[202]. Within the suggestive association category, we also replicated several previously associated risk genes; *PHACTR2*[203], *P2RX7*[204] and *PRKRA*[205]. Genetic variants in *PHACTR2 and PRKRA* were also reported by others to be associated to the risk of Parkinson's disease and ankylosing spondylitis, respectively[206,207]. Lately, the associated variants in *PRKRA* were retracted by the authors because they noticed that the alleged associated variants are rather a duplication of a true associations in the HLA region and in strong LD with them[208]. Nevertheless, we detected a frameshift mutation and this association survived the HLA alleles correction. The previously reported *P2RX7* association to MS was on the protective side[204] and this protective association was also observed in our study.

PPMS to RRMS comparisons have shown a number of significant associations even though the number of samples was limited. To RRMS, we identified an intergenic SNP on chromosome 4 that was significantly associated and an indel in *TRBV7-3* on chromosome 7 that was suggestively associated. *TRBV7-3* was reported to encode for the variable β chain of a T cell receptor specific for MBP[209,210]. To PPMS, we identified one significantly associated SNP in *FAM53C* and another six SNPs and one indel were suggestively associated. Four of the six SNPs are in the Lysine demethylases gene family (*KDM1B* and *KDM3B*). *FAM53C* encodes a protein that belongs to a family of nuclear proteins that are known to be involved in the neural tube development[211]. *KDM1B* and *KDM3B* proteins act as specific histone

demethylases leading to chromatin remodeling and similar histone modifications and chromatin remodeling have been highlighted in neuronal differentiation, survival and neurodegenerative diseases[212].

For identifying rare variants associations, we took a gene-based approach which resulted in observing multiple suggestive associations. The top two associations were *INO80D* to RRMS when compared to PBC and *GBP5* to PPMS when compared to RRMS. *GBP5* stimulates the inflammasome NRLP3 which is known to be tangled in the neuro-inflammatory process of many neurodegenerative diseases[213,214].

CNV analysis revealed 12 significant associations, including a deletion in *LCE3C* locus that had a lower frequency in PPMS (79%) compared to both RRMS (87%) and PBC (88%). Deletions in this locus is common in the general population and have also been associated to psoriasis and rheumatoid arthritis[215-217].

We mined for clinically significant mutations and a total of 187 and 31 clinically relevant pathogenic variants were identified in PPMS and RRMS, respectively. Focusing on mutations that were reported in MS and other neurological diseases, 15 were present in PPMS and 6 in RRMS. In PPMS, four mutations were hereditary spastic paraplegia (HSP) related mutations. This enrichment for HSP mutations in PPMS has also been reported recently by Jia et al[218]. Interestingly, another two were previous MS associated mutations, CYP27B1p.Arg389His and NR1H3p.Arg415Gln which is a subject of dilemma in regards of its association to MS[219-221].

These findings shed light on the genetic distinctiveness of PPMS from RRMS and on the fact that PPMS patients carry causative genetic variants of progressive neurological disorders that might help in understanding the progressive course of MS.

4.2 PAPER II

Based on the hypothesis that somatic mosaic genetic variants predispose specific immune cell to cross the BBB and become involved in the inflammatory process in the CNS we compared CNVs between CSF cells and PBMCs.

Using an Affymetrix array which is tailored for CNV detection we screened the whole genome for CNVs in paired CSF cells and PBMCs from six RRMS patients, one PPMS patient and two healthy controls. The comparison of these cells from either side of the BBB revealed deletions in the TCR loci; TRA/D, TRG and TRB in CSF cells. These deletions were confirmed using Taqman assays. When we sorted $CD4^+$, $CD8^+$ and γδ T cells from the PBMCs and compared them to the CSF cells a more complex pattern of these CNVs was detected. Deletions in TRG and TRA/D were consistent between all cell types except the TRA/D deletions in γδ T cells. Deletions in the TRB locus were mainly absent in γδ T cells and showed inter-individual and intra-variations between CSF, and $CD4^+$ and $CD8^+$ T cells. These findings suggested that these deletions are likely a result of the classic form of somatic mosaicism, TCR rearrangement. In a Japanese cohort, deletions in the TRG and TRA/D loci were associated with the risk of MS using DNA from PBMCs[222]. These deletions were validated only in T cells when different

white blood cells subsets were sorted to correct for the DNA source and concentration. Despite the deletions only being present in the sorted T cells, the authors speculated that their finding cannot be explained by TCR rearrangement as the complete J genes from the TRG loci were deleted. In another context, childhood allergic asthma, lower copies of the TRG loci was also associated to the risk of the disease[223]. This CNV in childhood allergic asthma was stated by the authors to be a mosaic somatic mutation as they were not inherited form their parents.

The absent TRA/D deletions in the γδ T cells suggests that the CSF cells are mainly αβ and not γδ T cells. Other studies have implicated γδ T cells in MS[33], expanded γδ T cells were shown to be present in acute MS plaques and CSF[30] and a subtype of γδ T cells was reported produce IFN- γ in a correlation with disease activity[224]. A fair comparison between isolated γδ T cells from the CSF and the periphery would be preferable, but this is challenging because of the small number of CSF cells. Even if we would manage to isolate γδ T cells from the CSF, another hurdle is to be able to extract enough yield of DNA. However, we did detect deletions in the TRB locus that are not expected according to the sequential TCR rearrangement[225]. Although this could be a result of sample contamination with αβ T cells, we dismiss that explanation as we then also expect to have detected deletions in the TRA locus, which we did not. Yet concurrent rearrangement of β, γ and δ loci[226] could explain these deletions.

The multifaceted deletions in the TRB locus have encouraged us to study and profile the T cell repertoire both in the periphery and central nervous system by sequencing the rearrangements in the TRB locus. Applying a frequency threshold of 0.1% for clonal expansion, the $CD4^+$ T cells were the most diverse compared to $CD8^+$ T cells and CSF cells, while $CD8^+$ T cells was the least diverse. In the periphery, it was also reported before that $CD8^+$ T cells are more expanded or less diverse than $CD4^+$ T cells[227]. A small number of CSF T cell clones were shared with the peripheral $CD4^+$and $CD8^+$ T cells, 7.2% and 4.9% of CSF T cell clones overlapped with $CD4^+$and $CD8^+$ T cells, respectively. This indicates a divergence in the TCR repertoire and compartmentalized expansion of the clones[228].

Dividing our TCR repertoire profiled cohort based on their clinical status of relapse (n=3) or remission (n=2) showed a less diverse or more expanded repertoire during relapse in all three compartments (CSF; p=0.0001, $CD4^+$ T cell; p=0.005 and $CD8^+$ T cell; p=0.037). This observation should be replicated in a larger cohort. Repertoire expansion or restriction might have a role in acute exacerbation of the disease. Agreeing with our results, a study with longitudinal CSF samples taken at relapse and remission, showed a reduced diversity of TRB repertoire at relapse[229]. In addition, natalizumab treatment reduced the repertoire expansion in the periphery but increased it in the CSF[230]. The after treatment increase in repertoire restriction may seem contradictory with our results, but as explained by the authors, this could be due to that the remaining cells in the CSF are selected T cells with autoantigen specificity[230].

It would be very interesting to know the target antigens of the identified clones. Therefore, we mined the antigen specific TCR sequences database, VDJdb[98], and identified the target antigens of several of the identified TCR sequences. The majority were viral antigens including the EBV antigens, which are of great interest because of EBV infection is a well-documented

environmental risk factor for MS[231]. In our identified repertoire, 3/4 of the expanded clones with known antigen specificity were targeting EBNA3A and EBNA3B antigens, but not EBNA1, which is the EBV antigens believed to be the main target for the humoral response[232].

In summary, with the ambition to identify novel variants that can be associated to MS and explain its missing heritability we screened for somatic CNVs and identified deletions in the TCR loci that were further revealed to be a consequence of specific TCR rearrangements which could potentially provide the basis for sub-populations of cells within the CNS. Our approach can assist in identifying autoreactive T cells and their corresponding antigens in MS and other autoimmune diseases.

4.3 PAPER III

After the successful GWASs in MS[127,128], a good number of genes have been highlighted for their possible role in the inflammatory process in MS. Here we studied how two of the effective immunomodulatory drugs, natalizumab and fingolimod, effect the peripheral protein expression levels of some of the genes with known involvement in the immune response[112,154,233,234]. The studied patients were first on natalizumab and then switched to fingolimod, as a precaution from developing PML after being positive for JC-virus antibodies[235].

We analysed changes in the levels of the soluble cytokine receptors; sIL-7Rα, sIL-2Rα, sIL-6R and sgp130, in serial plasma samples from before and after starting the treatment with natalizumab and then followed by fingolimod. Samples were collected approximately in the first 24-month treatment period for each drug and using the linear mixed effect model, we assessed the changes of the plasma levels of these soluble cytokine receptors from baseline and at 6, 12 and 24 months of treatment. This was done in two separate cohorts (cohort 1; n=47 and cohort 2; n=57) and similar results were obtained when analyzing each cohort separately or combined, therefore the results presented here is for the combined analysis. During natalizumab treatment the levels of sIL-7Rα and sgp130 decreased significantly while sIL-2Rα and sIL-6R levels did not change. The following fingolimod treatment lead to a significant increase of sIL-7Rα and sgp130, sIL-2Rα decreased significantly while sIL-6R remained unchanged.

Both natalizumab and fingolimod effect the peripheral lymphocyte count but in the opposite direction. Natalizumab increases the number of lymphocytes[236], hence we would expect an increase in the levels of these soluble receptors. Unpredictably, the only change we observed was a slight significant decrease in sgp130 and sIL-7Rα levels during treatment. Fingolimod decreases radically the number of lymphocytes in the periphery (70 % reduction) as a result of trapping them in the secondary lymphoid organs[57] and therefore a decrease in the levels of the soluble cytokine receptors is expected. Indeed, the levels of sIL-2Rα decreased significantly but it was disproportionate to the level of reduction in lymphocyte count. On the contrary, sgp130 and sIL-7Rα levels increased during fingolimod treatment and sIL-6R did not change. We checked for the correlation between the lymphocyte count and soluble cytokine receptors

levels and the only significant correlation was a negative correlation with sgp130, which could help in explaining the increase in the sgp130 during fingolimod treatment.

To assure that the changes in the levels of the soluble cytokine receptors is not a result of a cell count effect, we studied a group of patients treated with dimethyl fumarate (DMF). DMF causes lymphopenia through inducing apoptosis[237]. Our DMF treated cohort was divided into a group that developed lymphopenia and another with normal cell count. The results showed that both groups had no changes in the levels of the soluble cytokine receptors. The lymphopenia observed during fingolimod treatment is reversible as the cells are still present but just redistributed[238], hence one would expect a more sound effect of lymphopenia in DMF treatment if such an effect was the cause of the change of receptors levels during fingolimod treatment. Therefore, we hypothesized that the decrease of sIL-2Rα and increase of sgp130 and sIL-7Rα is a direct effect of fingolimod treatment. One reported effect is as an intranuclear histone deacetylase (HDAC) inhibitor, that enhances histone acetylation and acetylation and subsequently gene expression[239]. To test if the changes in the levels of these cytokine receptors is a direct effect of fingolimod and not a consequence of the sequestration of the cells, we cultured PBMCs for up to nine days in three different conditions; with fingolimod, phosphorylated fingolimod or just the vehicle. At day nine, the most significant results were that the levels of sIL-2Rα were significantly lower when the cells were treated with fingolimod than phosphorylated fingolimod or vehicle (Figure 6). These cell culture results may explain the decrease in the levels of sIL-2Rα that we observe in the plasma of fingolimod treated patients, but not the increase in sgp130 or sIL-7Rα.

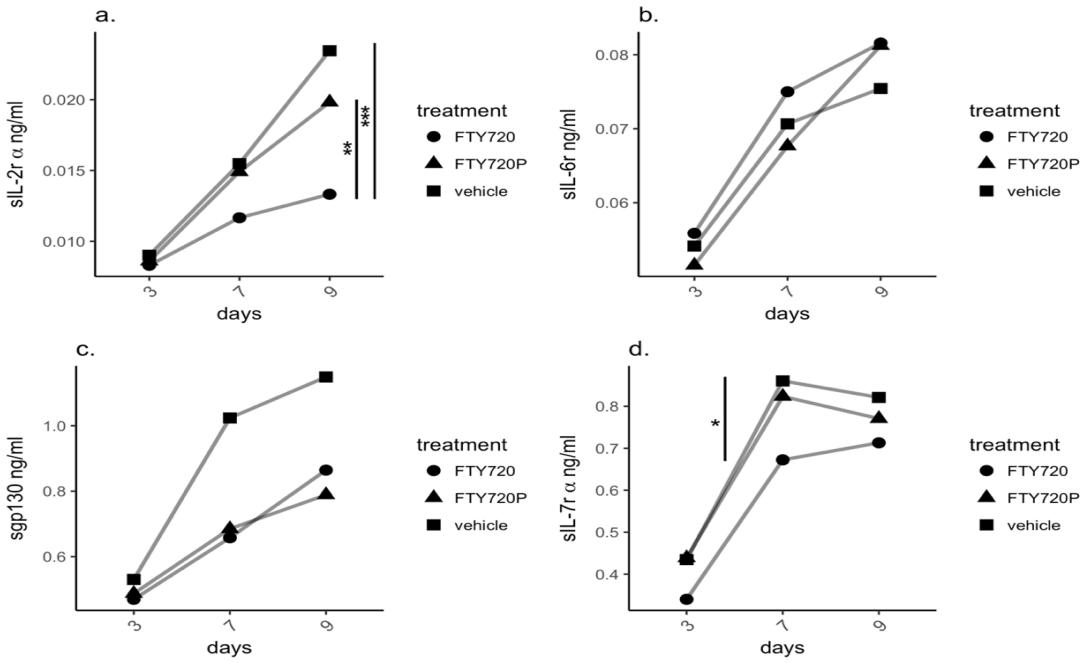

*Figure 6. PBMCs cultured for up to nine days in three conditions; with fingolimod or phosphorylated fingolimod or just the vehicle. The levels of a) sIL-2Rα b) sIL-6R c) sgp130 and d) sIL-7Rα were measured from the supernatant and the differences in the concentrations between the three conditions were tested for using the linear mixed model. * P value< 0.05, ** P value< 0.01, *** P value< 0.001.*

The increase in sgp130 during fingolimod treatment might result in a desired anti-inflammatory response. Sgp130 is known to block the pro-inflammatory IL-6 trans-signaling pathway[139] that otherwise favors the development of autoimmune inducing TH_{17}-cells[142,143] to inhibitory Tregs cells from naïve CD4 T cells[141]. TH_{17}-cells are involved in the neuro-inflammatory process[24,25] and reducing their number would be a favorable outcome. Fingolimod does reduce TH_{17}-cells count, however this effect has been attributed to their sequestration[240]. This might be plausible, nevertheless we suggest that blocking of their trans-signaling pathway through increasing sgp130 might contribute to reducing TH_{17}-cells.

We managed to replicate the association of the risk allele of the SNP rs6897932 with higher levels of sIL-7rα in plasma[112]. In addition, we observed a pharmacogenomic effect of the risk allele of rs6897932 during fingolimod treatment, in that sIL-7rα levels increase significantly more in patients homozygous for the risk allele than in other groups.

In this study, we were able to report the effect of the MS treatments on soluble cytokine receptors. These observed effects, even though they may appear as collateral or a consequence of the main effects, can still be valuable in other contexts, as they may have similar effects as anti-IL-6 or anti-IL2 drugs[241,242]. This encourages further studying of these treatments pharmacodynamics.

4.4 PAPER IV

With the general ultimate goal to tailor MS treatment for each patient, we took advantage of the advancement in the proteomics technology and utilized protein microarrays to identify potential biomarkers for MS treatment. We used the antibody suspension bead array, that had been useful in detecting biomarkers in neurological[243,244] and non-neurological disorders[245,246], and had proven to perform multiplex protein profiling in plasma samples[243,246].

Antibodies were selected for targeting and profiling proteins that are involved in the pathways of MS associated genetic risk variants[128], markers of inflammation, expressed in the CNS and had been interesting in previous MS protein profiling studies[244,247]. First, we profiled a total of 59 proteins in serial plasma samples of patients who have been on natalizumab and then switched to fingolimod (as a screening cohort we used the 44 patients of the paper III cohort 1). Here also we applied the linear mixed effect model to test for the changes in the measured protein levels during the first 24 months of each treatment period separately and a coefficient cut off of 0.25 and a p value < 0.001 were used. For a total of 10 antibodies, significant changes of the target protein levels were detected during the period of natalizumab treatment with no significant changes during fingolimod treatment (Figure 7).

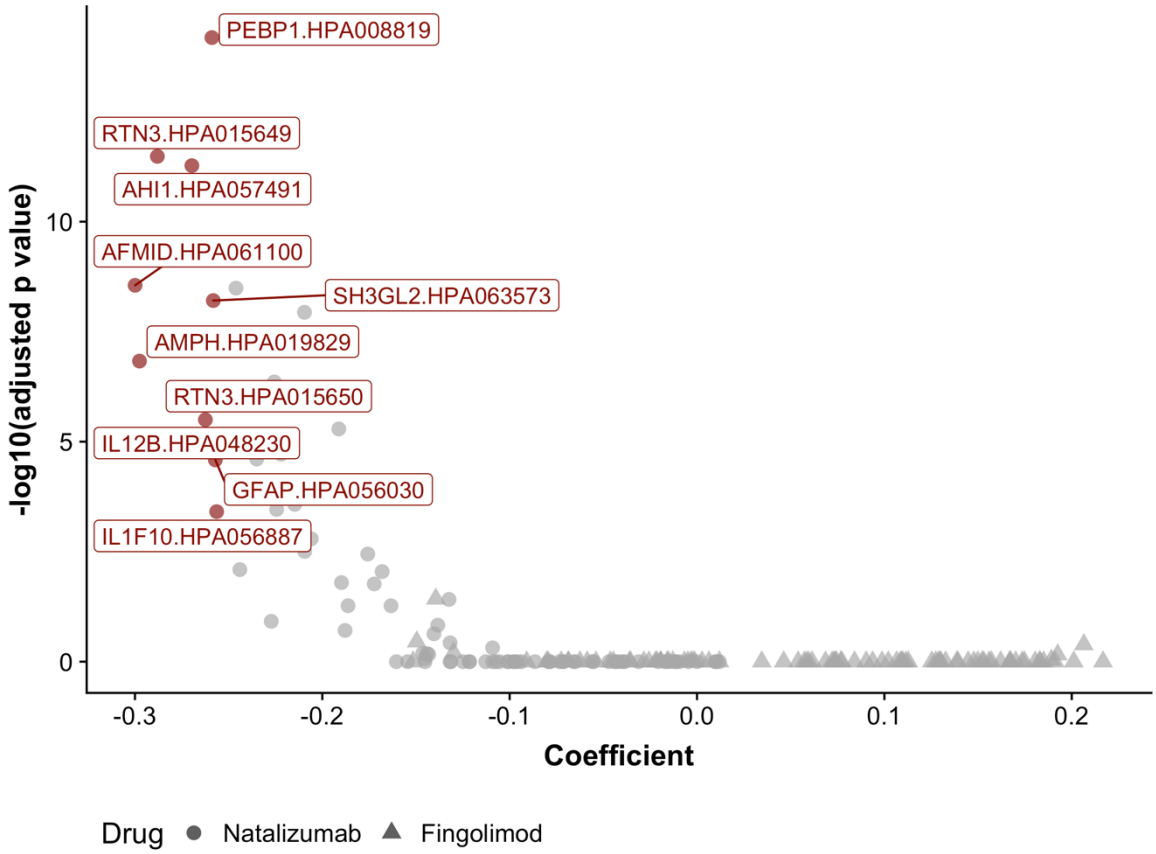

Figure 7. Scatter plot of the coefficients and adjusted p values obtained from the linear mixed effect model. Changes in the detected protein levels during natalizumab treatment (circles) and fingolimod (triangles) were considered significant (dark red) after using a cut off value of 0.25 for the coefficient and a p value < 0.001.

To confirm the reliability of the antibody array results we validated the specificity of the antibodies HPA008819 and HPA015649 targeting PEBP1 and RTN3, respectively. We developed two different types of in-house ELISA for each protein. For HPA008819, an in-house indirect ELISA and a sandwich ELISA and for RTN3 we developed an indirect ELISA and an inhibition ELISA. In the same screening cohort, we managed to reproduce the decrease in PEBP1 levels during natalizumab treatment and we also observed a decrease in the RTN3 levels, but they were not significant. Furthermore, in an independent cohort of MS patients treated with natalizumab, we replicated only the decrease in PEBP1 levels. One explanation for failing to reproduce the RTN3 observation from the protein array could be that, for the array, the plasma samples are heat-treated which affects the conformation of the protein and facilitate epitope accessibility[248]. However, we tested for this and observed no difference between heat-treated and non-treated plasma samples. Another possible explanation could be that the measurements obtained by the inhibition ELISA were not accurate.

Following the change in the disability outcomes, EDSS and MSSS, for all the patients included in this study reflects that the treatment response to natalizumab was quite good. To tell if the decrease in PEBP1 levels is a consequence of the positive treatment response or a treatment effect we combined both the screening and replication cohorts and studied the changes in

PEBP1 and RTN3 levels in two groups of patients, with the highest (moderate responders) or the lowest (mild responders) decrease values for the disability outcome, MSSS, during natalizumab treatment. Both PEBP1 and RTN3 decreased significantly in the moderate responder group while the levels did not change in the mild responders, which favors the treatment response rationalization. In support, the decline in PEBP1 was also observed in patients treated with DMF.

PEBP1 is expressed in the hippocampus and in patients with the neurodegenerative disease, Alzheimer's disease (AD), this protein is less present than in non-demented patients[171]. In addition, its derivative protein, HCNP, was reported to be high in the CSF of AD patients[249]. If we consider that the observed peripheral measurements of PEBP1 mirrors the neuro-inflammation and neurodegeneration, the decrease of PEBP1 during treatment might reflect the reduction in neurodegeneration.

In summary, we applied a semi hypothesis free approach to identify biomarkers for MS treatment using high-throughput proteomics technology. This approach succeeded in detecting the intra-individual change in the levels of several promising proteins, one of which is PEBP1, where we could reproduce and replicate its decline during treatment. This encourages further studies of the role of PEBP1 in MS and its potential as a biomarker for MS treatment.

5 CONCLUSIONS AND FUTURE PERSPECTIVES

Our knowledge of MS immuno-pathogenesis is increasing by the day especially with the current available immunomodulatory treatments (Figure 8). The scientific community had not really reached a consensus that MS is an autoimmune disease until these treatments have proven their efficiency.

There is a clear distinction between RRMS and PPMS clinically and pathologically to some extent, however genetically no variants have been identified to distinguish RRMS from PPMS[250]. This was reasoned by that the sufficient number of PPMS patients were not present, in other words, insufficient sample size. Nevertheless, a number of alleles have been suggested to be possibly associated with PPMS risk, such as APOE ε4 allele, *CASP8* GG homozygous SNP of rs2037815 and *IL-4R* Q551*R[251]. In addition, most of the identified MS associated gene variants fall outside the coding regions[82]. In paper I, we targeted the coding regions using whole exome sequencing in a modest number of PPMS and compared them to RRMS and PBC. One of the main findings is that in PPMS we didn't observe the strong protective association of class I allele HLA-A*02 for RRMS. In addition, we identified an indel in the TCR gene *TRBV7-3* suggestively associated to RRMS when compared to PPMS, this might emphasize the importance of antigen recognition and T cell activation in the course of RRMS but not in PPMS. Even though our PPMS cohort is of limited size, to our knowledge it is the largest cohort that an exome sequencing has been performed on. It would be interesting to see how increasing the size of the cohort will affect the suggestive associations, will they become stronger or disappear.

Most of the MS genetic studies are based on acquiring DNA from peripheral blood while the main target organs are found separated from the periphery by the BBB. This approach has nevertheless been successful in identifying germinal genetic variants under the common disease common variant hypothesis. On the other hand, there is an increasing interest in studying the role of somatic genetic variants in MS[252,253]. In paper II we studied the contribution of somatic variants to MS in a periphery - CNS comparison of peripheral and CSF cells, where peripheral cells were the reference samples. This approach, we believed, may detect genetic variants that will predispose immune cells to cross the BBB and take part in the inflammatory milieu in the CNS. We identified deletions in the TCR loci that are consistent with rearrangements of the TCR genes. This further emphasize the significance of the specificity of the immune reaction. The initial genome wide CNVs screening was performed in a small number of patients and healthy controls and in unsorted PBMCs and CSF cells. Adding more individuals and sorting the cells from the periphery and the CSF might help in revealing somatic variants in other regions. Nevertheless, it would be interesting if we could follow the TCR repertoire profile after treatment and isolate the expanded T cells and determine their targets using protein arrays.

The information of the associated MS risk genetic variants is available for the scientific community to build on and to understand how they contribute to the pathophysiology of MS. Functional studies have been successful in identifying the consequences of some of the associated variants[112,234]. In papers III and IV, we took advantage of these findings and studied the gene products of the MS associated loci and their potential use as biomarkers for monitoring natalizumab and fingolimod treatments. This was in conjunction of using longitudinally collected plasma samples which was possible because of the huge efforts in the longitudinal treatment surveillance cohorts. These cohorts provide the possibility of detecting intra-individual changes. In paper III, we observed changes during treatment in plasma levels of the soluble cytokine receptors sIL-7Rα, sIL-2Rα and sgp130 but not sIL-6r. Studying these changes in correlation to the disability measurements was quite challenging as most of the patients were good responders to the treatment. Whether these cytokine receptors could be used as biomarkers is a valid question that needs further studying, especially in a well characterized patient cohort on the basis of disease activity. The unchanged sIL-6R levels during treatment is puzzling even though both natalizumab and fingolimod have strong immunomodulatory effects and one would expect some changes to occur when either the immune cells are blocked from entering the CNS or captured in the lymph nodes. In addition to the gene products of the MS associated loci in paper IV we profiled the expression of CNS proteins and inflammatory markers in plasma where nine proteins exhibited decreasing levels during treatment. We choose to validate the top two proteins with the most significant changes, PEBP1 and RTN3. Yet, it would be interesting to validate the changes of the other seven proteins including e.g. GFAP, which previously has been reported as potential biomarker for disease progression[104], although its CSF levels did not change after natalizumb treatment[107].

If the identified genetic variants from the PPMS versus RRMS exome sequencing analysis are replicated they may support the question if PPMS is an independent entity from RRMS. In addition, they could also be candidates for disease progression genetic studies and assist the ongoing efforts in these studies. The approach that we took in paper II in identifying subpopulations of immune cells that cross into the CNS is in line with the efforts of finding therapies that target specific cell clones and not all cell lineage or subsets. This might help in reducing the possibility of infections and malignancies during treatment as the patient will not be immune compromised. The era of personalized medicine is here, medical practitioners are more aware now of the individual differences and they are adjusting to them, especially with the availability of multiple treatment options. In papers III and IV, we studied the effect of treatment on the protein expression in plasma with a main focus on the individual variation and observed variations in some of the proteins during treatment. Even though these variations did not correlate with the clinical variables, these observations might help in further understanding how the treatment works and to identify other possible effects on the immune system than the originally intended, which might assist in finding new indications for the treatments or the development of new treatments.

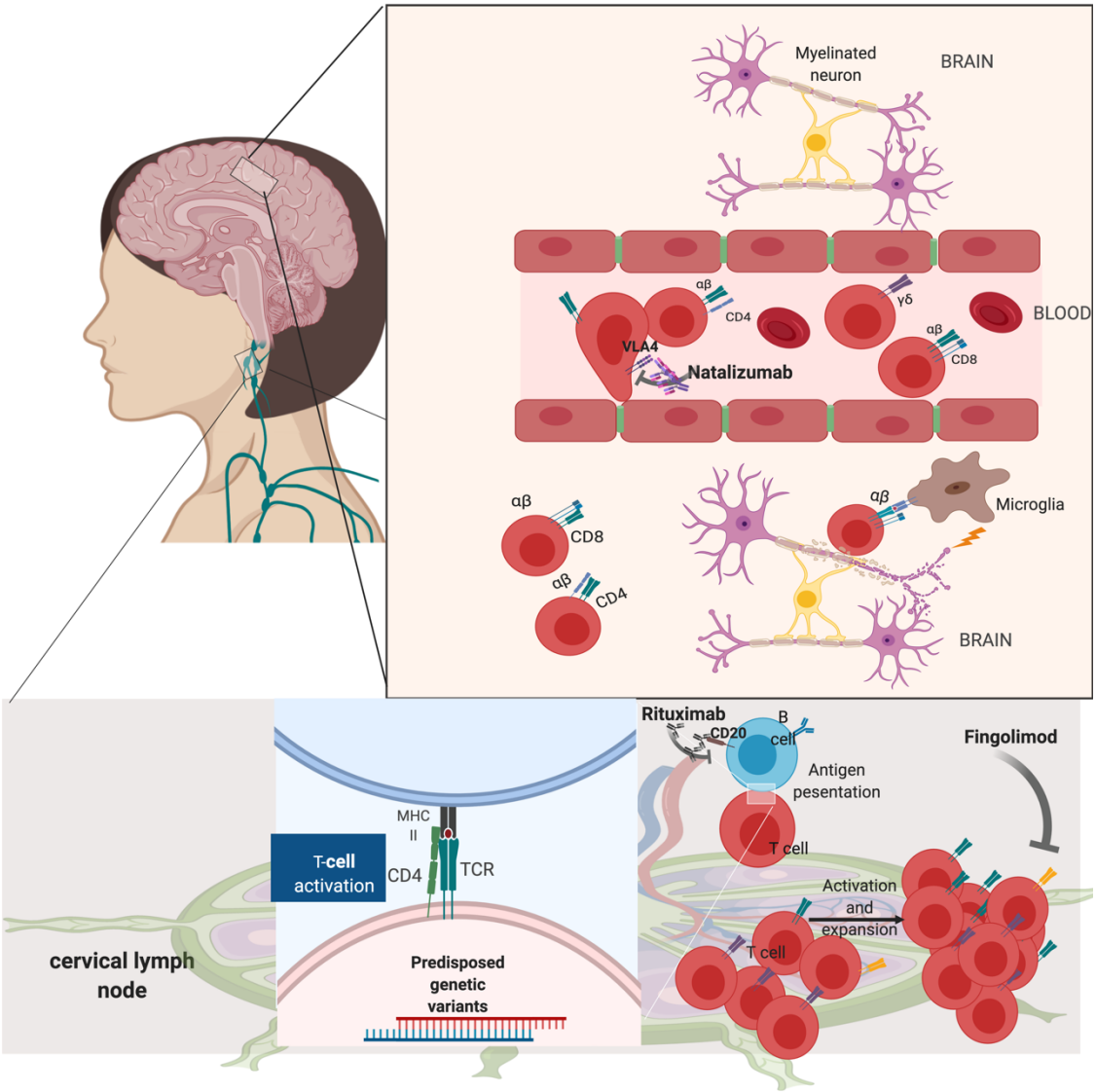

Figure 8. Proposed MS pathogenesis. *Created with biorender.com*

6 ACKNOWLEDGEMENTS

My main supervisor **Anna Glaser**, I would like to start by thanking you for your patience, maybe you should tell me what's your secret behind being always able to stay calm. we really worked hard to get this through and we managed at the end, I hope. I really appreciate that you were always there for support. Thank you also for the great long discussions that we had, scientifically or just about life outside the lab. I wish you all the best.

Jan Hillert, thank you for taking me into your group, it was really a pleasure to be part of it. You really provide for everyone in the group the fullest support. Thank you also for all the feedbacks and also for the nice scientific discussions we have in our group meetings, specially when you pick up the pen and start drawing graphs on the door.

Hans Grönlund, the man to go to when you are stuck with proteomics. Thank you for being a great supervisor, I like how you are enthusiastic about your research. We really had a long run with the HPA project but we made it.

Wangko Lundström, thank you for teaching me how to run ELISAs, that's what I have been doing for a very very long time. Thanks for helping also to start the soluble cytokines receptors project. All the best!

Anthony Wright, thank you for being my mentor and for your support.

Anna Fogdell-Hahn, for the great immunological discussions at the office and during journal clubs, thank you for making me read so many articles. It was really a learning period by itself during the journal clubs. Thank you for the support in the office and the lab.

Anna Mattsson, thank you for making sure that everything is there in the lab and in L8:00. You really make life very easy for everyone in the corridor. Vi alla uppskatta det fet mycket ☺

Ingegerd Löfving Arvholm, we really missed you in the lab after you started your pension and we always look forward to your visits. Thank you for all the support you gave me in the lab och för att jag fick öva svenska med dig på början när det var väldigt dåligt, nu det är bara så där svenska ☺.

Mohsen Khademi, you should have been one of my co-supervisor, but I knew that you were too busy, collecting all these huge cohorts that our research is built on, to supervise, that's why I never asked you about it. Thank you for all the help in my master period and then during my PhD.

Ingrid Kockum, thank you for taking me as a master student and for recommending me too Anna after that. I really enjoy the scientific atmosphere of your group meeting, thank you for letting me join them.

Tomas Olsson for all the cohorts.

Tojo James, it was great working with you on paper I, it will be a good one. Thank you for all the bioinformatics tips.

Ola B. Nilsson for the collaboration on paper IV, we really worked hard to get this project published.

Björn Evertsson, for the collaboration on paper II and all the samples phone calls.

Merja Kanerva for all the great help and support with sample collections in the lab at huddinge, I appreciate it. Kiitos! **Anja Finn**, for keeping everything under control in the lab. **Marjan Jahanpanah** for making me feel welcomed at Huddinge and for all the support with sample collections.

Peter Nilsson, **Anna Månberg**, **Carl Hamsten** and **Burcu Ayoglu** for your collaboration on paper IV.

Kyla Mckay, it was really nice having you as a newcomer in the group and also a good dancer when it's party time. **Ryan Ramanujam**, for all the movies critic and the gym tips, we should really finish the you, me and Ali competition. **Andrius Kavaliunas**, it was a pleasure being with you in the same group. **Virginija karrenbauer**, thank you for pushing to finish my Swedish studies and for all the collaborations from the clinic. **Katharina Fink**, we started together the last two studies and we finished them at the end, thank you for all the support from the clinic.

Leszek Stawiarz, for all the historical conversations and for your magnificent work with the MS registry. To all the previous Hillert's group members; **Helga Westerlind**, **Jenny Link**, **Kerstin Imrell**, **Stefanie Binzer** and **Omid Beiki.**

TID group, **Erik Holmgren**, **Guro Gafvelin**, **Claudia Querioz**, **Andreas Kaiser** and **Mattias Bronge** for sharing the cell culture room and your materials.

Christina Hermanrud, for being a great office mate for a long long time and for always having answers for lab and outside the lab. **Malin Ryner**, for assisting me with ELISAs when they failed and for being a nice person to be around. **Elin Engdahl**, for your enthusiasm for the environment. **Rasmus Gustafsson**, KIs no.1 runner, for welcoming when I moved to the group. **Nicky Dunn**, it was really nice to hang out when you are here and thanks for all the training motivations. **Natsya Kharlamova**, my new neighbour in the office, thank you for all the support during writing this thesis.

L8:00 floor-mates, present and former, I really had a great time in our floor. **Tomas Ekström** för alla de svenska ordspråk. **Atosa Estekizadeh**, **Joelle Rüegg**, **Louise Sjöholm**, **Malin Almgren, Oskar Karlsson, Xu Dawei**, **Hongya Han**, **Bingnan Li**, **Xiaolu Zhang**, **Yanxia Guo**, **Jingya Yu**, **Klas Strååt**, **Xiaotian Yuan**, **Xiangling Xing**, **Jenny Dahlström**, **Liu Tiantian**, **Yujiao Wu**, **Chuanyou Xia**, **Selina Parvin** and **Monica Ekberg**. KI gene; **Annika Eriksson, Katarina Gell**, **Malin Alvehus**, **Anna-Lee Jansén** and **Selim Sengul** for all your technical lab support and being kind**. Catharina Lavebratt**, **Martin Schalling**, **Laura Baqué**,

Linghua Kong, **Liu Yang**, **Yabin Wei, Anna Witasp, Ida Nilsson, Carlos Villaescusa, Miranda Stiernborg, Dzana Hukic, Björn Johansson, Judit Ozsvár, Jakob Schuy, Xinxia Chen, Iris Garcia, Paulina Luków** and **Annika Lindblom.**

Parvin, man I am going to miss your debates, even though they used to pop up randomly, at least I hope they improved my English, good luck with your post doc. **Paschalis Efstathopoulos**, it was really fun having you in the lab, sorry I missed your wedding while writing this thesis. **Mikael Ringh**, keep on smiling and shining, we went together through all PhD up and downs, but now we are almost done. Thank you for all the help with the figure designs. **Vincent, Don V or professor V**, it's hard to tell which one you will become, you gather people around you for EU causes in the weekends and in the weekdays you switch them with students for the R club. Soon we will have a first authorship paper together don't give up. Guys I'm going to miss our chats at the kitchen.

Friends from the 4th and 5th floor, **Shahin Aeinehband, Joan Rouf, Sohel Mia, Mélanie Thessén Hedreul, Henna Salo, Sunjay Fernandes** and **Niyaz Yoosuf** always fun to be with you guys. **Venus Azhary** and **Brinda Dev** for your support during my master. **Samina Asad, Pernilla Strid, Alexandra Gyllenberg, Jesse Huang, Katarina Tengvall, Tenghao Zheng, Xingrong Liu** and **Thomas Moridi** for the nice discussions in Ingrids meetings.

CNS adminstration, **Kseniya Hartvigsson** for making all our PhD documents are right in time and being a gym buddy, **Gullan Rydén** for always making sure our numbers are right, **Joakim Airaxin** for contributing to science, **Åsa Malmquist** for helping with all the paper work. **Elin Johansson,** for being my first HR personal to work with, **Elin Hammergård** for taking over from the other **Elin**, **Anders Ihrfors**, **Karin Blomberg** and **Alexandre Shirasaka Beckman** always nice chatting with you.

IT guys, **Daniel Uvehag**, **Christer Uvehag Christian Dalin, Timmy Hellström** and **Jobel Teklebrhan**, thank you for your support.

To the IMSE project managers **Linda Forsberg** and **Susanne Johansson**.

Ali Manouchehrinia, my brother at work and outside work and bro in the gym. Thank you for jumping in with the stats in the last two papers, that was essential input to both studies. Keep up with the hard work, I see you soon having your own group.

Magdalena Lindén, the angel of CMM. I am really lucky that you supervised me during my master project, I learned a lot from you and you helped me a lot. I am really glad that part of our work is now published.

Nada Omer, my sister from another mother and dancing partner since we were kids. This all began from you helping me to get a master project, thank you! Thank you for all the social support when I started working in CMM. I Wish you all the best in the clinic.

Amir Saeed, my brother from another mother, for sharing you room with a stranger when I moved to Stockholm and for your patience. Thank you for the endless unquestionable help and support during my first days here and after.

To friends who worked at different departments, **Eva Hagel, Waleed Tajeddinn, Radwa Almamoun, Ahmed Abdelwahid, Osman Ahmed, Mustafa Ismail, Amira Elkharusi, Husameldin Elnour, Fahd Al-Zadjali, Mohamed Ali, Ahmed Omer, Tagrid Salih, and Abubakr Obeidallah.**

Ahmed Osman, for always listening to my complains during lunch and our way back from the mosque. **Waleed Elamin** for the nice talks during lunch.

To family and friends close by **Toota, Berhan**, **Jasmin Kadir** (when you come from Copenhagen), **Aymen Mahamed**, **Balsam Haseeb**, **Amani Abdelkhalig**, **Kamal Yassin, Amel, Abir Ali, Fowzia (Fo Fo)**, **Ali Eltom, Adila Elfatih, Mohamed Salah, Siddig Elmakashfi** and **Babiker Elobeid** and his family.

To **Misara Sanosi**, **Mohamed Ali** and **Mohmed Musa** for making Husby feels like home.

Ibrahim Bedri, my uncle, it was you who suggested a master in Sweden, thank you for all the support during my stay and even before that when you got me the winter clothing to be ready before I set my foot here.

Boran Bedri, Gasim Bedri, Elnigoumi Bedri and all the family in Khartoum.

Sadia Ahmed, for supporting me, Nejat and Mae, thank you.

To my great sisters **Shams** and **Shahinaz** and their beautiful families, thank you for your support before coming to Sweden and during my stay here.

To my mother, **Sulafa**, no words can be enough to thank you for your unconditional love and care.

To my wife **Nejat**, I came to study in Sweden but then I also got married in Sweden, love happens, thank you for understanding the hard times that I had, for your patience, for the "fine dining" at home and for giving us the fruit of our love Mae ❤.

Mae, now we can play, play and play ;)

8 REFERENCES

1 Compston, A. & Coles, A. Multiple sclerosis. *Lancet* **359**, 1221-1231, doi:10.1016/s0140-6736(02)08220-x (2002).

2 Murray, T. J. The history of multiple sclerosis: the changing frame of the disease over the centuries. *J Neurol Sci* **277 Suppl 1**, S3-8, doi:10.1016/s0022-510x(09)70003-6 (2009).

3 Pearce, J. M. Historical descriptions of multiple sclerosis. *Eur Neurol* **54**, 49-53, doi:10.1159/000087387 (2005).

4 Milo, R. & Kahana, E. Multiple sclerosis: geoepidemiology, genetics and the environment. *Autoimmun Rev* **9**, A387-394, doi:10.1016/j.autrev.2009.11.010 (2010).

5 Simpson, S., Jr., Blizzard, L., Otahal, P., Van der Mei, I. & Taylor, B. Latitude is significantly associated with the prevalence of multiple sclerosis: a meta-analysis. *J Neurol Neurosurg Psychiatry* **82**, 1132-1141, doi:10.1136/jnnp.2011.240432 (2011).

6 Ahlgren, C., Oden, A. & Lycke, J. High nationwide prevalence of multiple sclerosis in Sweden. *Mult Scler* **17**, 901-908, doi:10.1177/1352458511403794 (2011).

7 Miller, D. H. & Leary, S. M. Primary-progressive multiple sclerosis. *Lancet Neurol* **6**, 903-912, doi:10.1016/s1474-4422(07)70243-0 (2007).

8 McDonald, W. I. *et al.* Recommended diagnostic criteria for multiple sclerosis: guidelines from the International Panel on the diagnosis of multiple sclerosis. *Ann Neurol* **50**, 121-127 (2001).

9 Polman, C. H. *et al.* Diagnostic criteria for multiple sclerosis: 2010 revisions to the McDonald criteria. *Ann Neurol* **69**, 292-302, doi:10.1002/ana.22366 (2011).

10 Thompson, A. J. *et al.* Diagnosis of multiple sclerosis: 2017 revisions of the McDonald criteria. *Lancet Neurol* **17**, 162-173, doi:10.1016/S1474-4422(17)30470-2 (2018).

11 Kurtzke, J. F. Rating neurologic impairment in multiple sclerosis: an expanded disability status scale (EDSS). *Neurology* **33**, 1444-1452 (1983).

12 Noseworthy, J. H., Lucchinetti, C., Rodriguez, M. & Weinshenker, B. G. Multiple sclerosis. *N Engl J Med* **343**, 938-952, doi:10.1056/nejm200009283431307 (2000).

13 Hafler, D. A. Multiple sclerosis. *J Clin Invest* **113**, 788-794, doi:10.1172/jci21357 (2004).

14 Kutzelnigg, A. & Lassmann, H. Pathology of multiple sclerosis and related inflammatory demyelinating diseases. *Handb Clin Neurol* **122**, 15-58, doi:10.1016/b978-0-444-52001-2.00002-9 (2014).

15 Sospedra, M. & Martin, R. Immunology of Multiple Sclerosis. *Semin Neurol* **36**, 115-127, doi:10.1055/s-0036-1579739 (2016).

16 Hauser, S. L. *et al.* Immunohistochemical analysis of the cellular infiltrate in multiple sclerosis lesions. *Ann Neurol* **19**, 578-587, doi:10.1002/ana.410190610 (1986).

17 Booss, J., Esiri, M. M., Tourtellotte, W. W. & Mason, D. Y. Immunohistological analysis of T lymphocyte subsets in the central nervous system in chronic progressive multiple sclerosis. *J Neurol Sci* **62**, 219-232 (1983).

18 Babbe, H. *et al.* Clonal expansions of CD8(+) T cells dominate the T cell infiltrate in active multiple sclerosis lesions as shown by micromanipulation and single cell polymerase chain reaction. *J Exp Med* **192**, 393-404 (2000).

19 Fogdell-Hahn, A., Ligers, A., Gronning, M., Hillert, J. & Olerup, O. Multiple sclerosis: a modifying influence of HLA class I genes in an HLA class II associated autoimmune disease. *Tissue Antigens* **55**, 140-148 (2000).

20 Friese, M. A. *et al.* Opposing effects of HLA class I molecules in tuning autoreactive CD8+ T cells in multiple sclerosis. *Nat Med* **14**, 1227-1235, doi:10.1038/nm.1881 (2008).

21 Murphy, K. M. & Reiner, S. L. The lineage decisions of helper T cells. *Nature Reviews Immunology* **2**, 933-944, doi:doi:10.1038/nri954 (2002).

22 Harrington, L. E. *et al.* Interleukin 17-producing CD4+ effector T cells develop via a lineage distinct from the T helper type 1 and 2 lineages. *Nat Immunol* **6**, 1123-1132, doi:10.1038/ni1254 (2005).

23 Lassmann, H. & Ransohoff, R. M. The CD4-Th1 model for multiple sclerosis: a critical [correction of crucial] re-appraisal. *Trends Immunol* **25**, 132-137, doi:10.1016/j.it.2004.01.007 (2004).

24 Kebir, H. *et al.* Human TH17 lymphocytes promote blood-brain barrier disruption and central nervous system inflammation. *Nat Med* **13**, 1173-1175, doi:10.1038/nm1651 (2007).

25 Brucklacher-Waldert, V., Stuerner, K., Kolster, M., Wolthausen, J. & Tolosa, E. Phenotypical and functional characterization of T helper 17 cells in multiple sclerosis. *Brain* **132**, 3329-3341, doi:10.1093/brain/awp289 (2009).

26 Tzartos, J. S. *et al.* Interleukin-17 production in central nervous system-infiltrating T cells and glial cells is associated with active disease in multiple sclerosis. *Am J Pathol* **172**, 146-155, doi:10.2353/ajpath.2008.070690 (2008).

27 Muranski, P. & Restifo, N. P. Essentials of Th17 cell commitment and plasticity. *Blood* **121**, 2402-2414, doi:10.1182/blood-2012-09-378653 (2013).

28 Murphy, K. M. & Stockinger, B. Effector T cell plasticity: flexibility in the face of changing circumstances. *Nat Immunol* **11**, 674-680, doi:10.1038/ni.1899 (2010).

29 Carding, S. R. & Egan, P. J. γδ T cells: functional plasticity and heterogeneity. *Nature Reviews Immunology* **2**, 336, doi:10.1038/nri797 (2002).

30 Van Kaer, L., Postoak, J. L., Wang, C., Yang, G. & Wu, L. Innate, innate-like and adaptive lymphocytes in the pathogenesis of MS and EAE. *Cellular & Molecular Immunology*, doi:10.1038/s41423-019-0221-5 (2019).

31 Ribot, J. C. *et al.* CD27 is a thymic determinant of the balance between interferon-gamma- and interleukin 17-producing gammadelta T cell subsets. *Nature immunology* **10**, 427-436, doi:10.1038/ni.1717 (2009).

32 Singh, A. K. *et al.* High Interferon-γ Uniquely in Vδ1 T Cells Correlates with Markers of Inflammation and Axonal Damage in Early Multiple Sclerosis. *Frontiers in Immunology* **8**, 260 (2017).

33 Schirmer, L., Rothhammer, V., Hemmer, B. & Korn, T. Enriched CD161high CCR6+ γδ T Cells in the Cerebrospinal Fluid of Patients With Multiple SclerosisCD161high

CCR6+ γδ T Cells in Multiple Sclerosis. *JAMA Neurology* **70**, 345-351, doi:10.1001/2013.jamaneurol.409 (2013).

34 Serafini, B., Rosicarelli, B., Magliozzi, R., Stigliano, E. & Aloisi, F. Detection of ectopic B-cell follicles with germinal centers in the meninges of patients with secondary progressive multiple sclerosis. *Brain Pathol* **14**, 164-174 (2004).

35 Hauser, S. L. *et al.* B-cell depletion with rituximab in relapsing-remitting multiple sclerosis. *N Engl J Med* **358**, 676-688, doi:10.1056/NEJMoa0706383 (2008).

36 Kappos, L. *et al.* Ocrelizumab in relapsing-remitting multiple sclerosis: a phase 2, randomised, placebo-controlled, multicentre trial. *Lancet* **378**, 1779-1787, doi:10.1016/S0140-6736(11)61649-8 (2011).

37 Cross, A. H., Stark, J. L., Lauber, J., Ramsbottom, M. J. & Lyons, J. A. Rituximab reduces B cells and T cells in cerebrospinal fluid of multiple sclerosis patients. *J Neuroimmunol* **180**, 63-70, doi:10.1016/j.jneuroim.2006.06.029 (2006).

38 Li, R., Patterson, K. R. & Bar-Or, A. Reassessing B cell contributions in multiple sclerosis. *Nature Immunology* **19**, 696-707, doi:10.1038/s41590-018-0135-x (2018).

39 Jacobs, L. D. *et al.* Intramuscular interferon beta-1a for disease progression in relapsing multiple sclerosis. The Multiple Sclerosis Collaborative Research Group (MSCRG). *Ann Neurol* **39**, 285-294, doi:10.1002/ana.410390304 (1996).

40 Panitch, H., Miller, A., Paty, D. & Weinshenker, B. Interferon beta-1b in secondary progressive MS: results from a 3-year controlled study. *Neurology* **63**, 1788-1795 (2004).

41 Leary, S. M. *et al.* Interferon beta-1a in primary progressive MS: an exploratory, randomized, controlled trial. *Neurology* **60**, 44-51 (2003).

42 Dhib-Jalbut, S. & Marks, S. Interferon-beta mechanisms of action in multiple sclerosis. *Neurology* **74 Suppl 1**, S17-24, doi:10.1212/WNL.0b013e3181c97d99 (2010).

43 Arnon, R. & Sela, M. Immunomodulation by the copolymer glatiramer acetate. *J Mol Recognit* **16**, 412-421, doi:10.1002/jmr.628 (2003).

44 Johnson, K. P. *et al.* Copolymer 1 reduces relapse rate and improves disability in relapsing-remitting multiple sclerosis: results of a phase III multicenter, double-blind placebo-controlled trial. The Copolymer 1 Multiple Sclerosis Study Group. *Neurology* **45**, 1268-1276 (1995).

45 Comi, G., Filippi, M. & Wolinsky, J. S. European/Canadian multicenter, double-blind, randomized, placebo-controlled study of the effects of glatiramer acetate on magnetic resonance imaging--measured disease activity and burden in patients with relapsing multiple sclerosis. European/Canadian Glatiramer Acetate Study Group. *Ann Neurol* **49**, 290-297 (2001).

46 Papadopoulou, A., Kappos, L. & Sprenger, T. Teriflunomide for oral therapy in multiple sclerosis. *Expert Rev Clin Pharmacol* **5**, 617-628, doi:10.1586/ecp.12.56 (2012).

47 Vermersch, P. *et al.* Teriflunomide versus subcutaneous interferon beta-1a in patients with relapsing multiple sclerosis: a randomised, controlled phase 3 trial. *Mult Scler* **20**, 705-716, doi:10.1177/1352458513507821 (2014).

48 Gold, R. *et al.* Placebo-Controlled Phase 3 Study of Oral BG-12 for Relapsing Multiple Sclerosis. http://dx.doi.org/10.1056/NEJMoa1114287, doi:NJ201209203671206 (2012).

49 Linker, R. A. *et al.* Fumaric acid esters exert neuroprotective effects in neuroinflammation via activation of the Nrf2 antioxidant pathway. *Brain* **134**, 678-692, doi:10.1093/brain/awq386 (2011).

50 Schulze-Topphoff, U. *et al.* Dimethyl fumarate treatment induces adaptive and innate immune modulation independent of Nrf2. *Proc Natl Acad Sci U S A* **113**, 4777-4782, doi:10.1073/pnas.1603907113 (2016).

51 Carlström, K. E. *et al.* Therapeutic efficacy of dimethyl fumarate in relapsing-remitting multiple sclerosis associates with ROS pathway in monocytes. *Nature Communications* **10**, 3081, doi:10.1038/s41467-019-11139-3 (2019).

52 Polman, C. H. *et al.* A randomized, placebo-controlled trial of natalizumab for relapsing multiple sclerosis. *N Engl J Med* **354**, 899-910, doi:10.1056/NEJMoa044397 (2006).

53 Clifford, D. B. *et al.* Natalizumab-associated progressive multifocal leukoencephalopathy in patients with multiple sclerosis: lessons from 28 cases. *Lancet Neurol* **9**, 438-446, doi:10.1016/s1474-4422(10)70028-4 (2010).

54 Warnke, C. *et al.* Changes to anti-JCV antibody levels in a Swedish national MS cohort. *J Neurol Neurosurg Psychiatry* **84**, 1199-1205, doi:10.1136/jnnp-2012-304332 (2013).

55 Brinkmann, V. *et al.* The immune modulator FTY720 targets sphingosine 1-phosphate receptors. *J Biol Chem* **277**, 21453-21457, doi:10.1074/jbc.C200176200 (2002).

56 Calabresi, P. A. *et al.* Safety and efficacy of fingolimod in patients with relapsing-remitting multiple sclerosis (FREEDOMS II): a double-blind, randomised, placebo-controlled, phase 3 trial. *Lancet Neurol* **13**, 545-556, doi:10.1016/s1474-4422(14)70049-3 (2014).

57 Kappos, L. *et al.* A placebo-controlled trial of oral fingolimod in relapsing multiple sclerosis. *N Engl J Med* **362**, 387-401, doi:10.1056/NEJMoa0909494 (2010).

58 Tyler, K. L. in *JAMA Neurol* Vol. 72 10-13 (2015).

59 Reff, M. E. *et al.* Depletion of B cells in vivo by a chimeric mouse human monoclonal antibody to CD20. *Blood* **83**, 435-445 (1994).

60 Berntsson, S. G. *et al.* Rapidly increasing off-label use of rituximab in multiple sclerosis in Sweden - Outlier or predecessor? *Acta Neurol Scand* **138**, 327-331, doi:10.1111/ane.12963 (2018).

61 Sorensen, P. S. & Sellebjerg, F. Pulsed immune reconstitution therapy in multiple sclerosis. *Ther Adv Neurol Disord* **12**, 1756286419836913, doi:10.1177/1756286419836913 (2019).

62 Muraro, P. A. *et al.* Autologous haematopoietic stem cell transplantation for treatment of multiple sclerosis. *Nature Reviews Neurology* **13**, 391, doi:10.1038/nrneurol.2017.81 (2017).

63 Faissner, S., Plemel, J. R., Gold, R. & Yong, V. W. Progressive multiple sclerosis: from pathophysiology to therapeutic strategies. *Nature Reviews Drug Discovery*, doi:10.1038/s41573-019-0035-2 (2019).

64 Montalban, X. *et al.* Ocrelizumab versus Placebo in Primary Progressive Multiple Sclerosis. *N Engl J Med* **376**, 209-220, doi:10.1056/NEJMoa1606468 (2017).

65 Kappos, L. *et al.* Siponimod versus placebo in secondary progressive multiple sclerosis (EXPAND): a double-blind, randomised, phase 3 study. *Lancet* **391**, 1263-1273, doi:10.1016/S0140-6736(18)30475-6 (2018).

66 Rothman, K. *Epidemiology: an introduction*. (Oxford University Press, 2002).

67 Ascherio, A. & Munger, K. L. Epidemiology of Multiple Sclerosis: From Risk Factors to Prevention-An Update. *Semin Neurol* **36**, 103-114, doi:10.1055/s-0036-1579693 (2016).

68 Munger, K. L., Levin, L. I., Hollis, B. W., Howard, N. S. & Ascherio, A. Serum 25-hydroxyvitamin D levels and risk of multiple sclerosis. *Jama* **296**, 2832-2838, doi:10.1001/jama.296.23.2832 (2006).

69 Ascherio, A. *et al.* Vitamin D as an early predictor of multiple sclerosis activity and progression. *JAMA Neurol* **71**, 306-314, doi:10.1001/jamaneurol.2013.5993 (2014).

70 Hedstrom, A. K., Baarnhielm, M., Olsson, T. & Alfredsson, L. Tobacco smoking, but not Swedish snuff use, increases the risk of multiple sclerosis. *Neurology* **73**, 696-701, doi:10.1212/WNL.0b013e3181b59c40 (2009).

71 Ramanujam, R. *et al.* Effect of Smoking Cessation on Multiple Sclerosis Prognosis. *JAMA Neurol* **72**, 1117-1123, doi:10.1001/jamaneurol.2015.1788 (2015).

72 Hedstrom, A. K., Akerstedt, T., Hillert, J., Olsson, T. & Alfredsson, L. Shift work at young age is associated with increased risk for multiple sclerosis. *Ann Neurol* **70**, 733-741, doi:10.1002/ana.22597 (2011).

73 Thacker, E. L., Mirzaei, F. & Ascherio, A. Infectious mononucleosis and risk for multiple sclerosis: a meta-analysis. *Ann Neurol* **59**, 499-503, doi:10.1002/ana.20820 (2006).

74 Sadovnick, A. D. & Baird, P. A. THE FAMILIAL NATURE OF MULTIPLE-SCLEROSIS - AGE-CORRECTED EMPIRIC RECURRENCE RISKS FOR CHILDREN AND SIBLINGS OF PATIENTS. *Neurology* **38**, 990-991 (1988).

75 Ebers, G. C. *et al.* A POPULATION-BASED STUDY OF MULTIPLE-SCLEROSIS IN TWINS. *New England Journal of Medicine* **315**, 1638-1642, doi:10.1056/nejm198612253152603 (1986).

76 O'Gorman, C., Lin, R., Stankovich, J. & Broadley, S. A. Modelling genetic susceptibility to multiple sclerosis with family data. *Neuroepidemiology* **40**, 1-12, doi:10.1159/000341902 (2013).

77 Westerlind, H. *et al.* Modest familial risks for multiple sclerosis: a registry-based study of the population of Sweden. *Brain* **137**, 770-778, doi:10.1093/brain/awt356 (2014).

78 Jersild, C. *et al.* Histocompatibility determinants in multiple sclerosis, with special reference to clinical course. *Lancet* **2**, 1221-1225 (1973).

79 Olerup, O. & Hillert, J. HLA class II-associated genetic susceptibility in multiple sclerosis: a critical evaluation. *Tissue Antigens* **38**, 1-15 (1991).

80 Weber, F. *et al.* IL2RA and IL7RA genes confer susceptibility for multiple sclerosis in two independent European populations. *Genes Immun* **9**, 259-263, doi:10.1038/gene.2008.14 (2008).

81 Lundmark, F. *et al.* Variation in interleukin 7 receptor alpha chain (IL7R) influences risk of multiple sclerosis. *Nat Genet* **39**, 1108-1113, doi:10.1038/ng2106 (2007).

82 Baranzini, S. E. & Oksenberg, J. R. The Genetics of Multiple Sclerosis: From 0 to 200 in 50 Years. *Trends Genet* **33**, 960-970, doi:10.1016/j.tig.2017.09.004 (2017).

83 Lill, C. M. Recent advances and future challenges in the genetics of multiple sclerosis. *Front Neurol* **5**, 130, doi:10.3389/fneur.2014.00130 (2014).

84 Manolio, T. A. *et al.* Finding the missing heritability of complex diseases. *Nature* **461**, 747-753, doi:10.1038/nature08494 (2009).

85 Womack, J. E., Jang, H. J. & Lee, M. O. Genomics of complex traits. *Ann N Y Acad Sci* **1271**, 33-36, doi:10.1111/j.1749-6632.2012.06733.x (2012).

86 Strachan, T. & Read, A. P. *Human molecular genetics*. 4th ed. edn, (Garland Science, 2010).

87 Youssoufian, H. & Pyeritz, R. E. Mechanisms and consequences of somatic mosaicism in humans. *Nat Rev Genet* **3**, 748-758, doi:10.1038/nrg906 (2002).

88 Piotrowski, A. *et al.* Somatic mosaicism for copy number variation in differentiated human tissues. *Hum Mutat* **29**, 1118-1124, doi:10.1002/humu.20815 (2008).

89 O'Huallachain, M., Karczewski, K. J., Weissman, S. M., Urban, A. E. & Snyder, M. P. Extensive genetic variation in somatic human tissues. *Proc Natl Acad Sci U S A* **109**, 18018-18023, doi:10.1073/pnas.1213736109 (2012).

90 Hodgkin, P. D., Heath, W. R. & Baxter, A. G. The clonal selection theory: 50 years since the revolution. *Nat Immunol* **8**, 1019-1026, doi:10.1038/ni1007-1019 (2007).

91 Kim, J. *et al.* Somatic deletions implicated in functional diversity of brain cells of individuals with schizophrenia and unaffected controls. *Sci Rep* **4**, 3807, doi:10.1038/srep03807 (2014).

92 Pamphlett, R., Morahan, J. M., Luquin, N. & Yu, B. Looking for differences in copy number between blood and brain in sporadic amyotrophic lateral sclerosis. *Muscle Nerve* **44**, 492-498, doi:10.1002/mus.22095 (2011).

93 Abbas, A. K., Lichtman, A. H. & Pillai, S. *Cellular and molecular immunology*. Eighth edition. edn, (Elsevier Saunders, 2015).

94 Matos, T. R., de Rie, M. A. & Teunissen, M. B. M. Research Techniques Made Simple: High-Throughput Sequencing of the T-Cell Receptor. *J Invest Dermatol* **137**, e131-e138, doi:10.1016/j.jid.2017.04.001 (2017).

95 Robins, H. S. *et al.* Comprehensive assessment of T-cell receptor beta-chain diversity in alphabeta T cells. *Blood* **114**, 4099-4107, doi:10.1182/blood-2009-04-217604 (2009).

96 De Simone, M., Rossetti, G. & Pagani, M. Single Cell T Cell Receptor Sequencing: Techniques and Future Challenges. *Frontiers in Immunology* **9**, doi:10.3389/fimmu.2018.01638 (2018).

97 Tickotsky, N., Sagiv, T., Prilusky, J., Shifrut, E. & Friedman, N. McPAS-TCR: a manually curated catalogue of pathology-associated T cell receptor sequences. *Bioinformatics* **33**, 2924-2929, doi:10.1093/bioinformatics/btx286 (2017).

98 Shugay, M. *et al.* VDJdb: a curated database of T-cell receptor sequences with known antigen specificity. *Nucleic Acids Res* **46**, D419-D427, doi:10.1093/nar/gkx760 (2018).

99 Planas, R., Metz, I., Martin, R. & Sospedra, M. Detailed Characterization of T Cell Receptor Repertoires in Multiple Sclerosis Brain Lesions. *Front Immunol* **9**, 509, doi:10.3389/fimmu.2018.00509 (2018).

100 Group., B. D. W. Biomarkers and surrogate endpoints: preferred definitions and conceptual framework. *Clin Pharmacol Ther* **69**, 89-95, doi:10.1067/mcp.2001.113989 (2001).

101 Lowenthal, A., Vansande, M. & Karcher, D. The differential diagnosis of neurological diseases by fractionating electrophoretically the CSF gamma-globulins. *J Neurochem* **6**, 51-56 (1960).

102 Arrambide, G. *et al.* The value of oligoclonal bands in the multiple sclerosis diagnostic criteria. *Brain* **141**, 1075-1084, doi:10.1093/brain/awy006 (2018).

103 Zivadinov, R. & Leist, T. P. Clinical-magnetic resonance imaging correlations in multiple sclerosis. *J Neuroimaging* **15**, 10s-21s, doi:10.1177/1051228405283291 (2005).

104 Malmeström, C., Haghighi, S., Rosengren, L., Andersen, O. & Lycke, J. Neurofilament light protein and glial fibrillary acidic protein as biological markers in MS. *Neurology* **61**, 1720-1725, doi:10.1212/01.wnl.0000098880.19793.b6 (2003).

105 Salzer, J., Svenningsson, A. & Sundström, P. Neurofilament light as a prognostic marker in multiple sclerosis. doi:10.1177/1352458509359725 (2010).

106 Kuhle, J. *et al.* Serum neurofilament light chain in early relapsing remitting MS is increased and correlates with CSF levels and with MRI measures of disease severity. *Mult Scler*, doi:10.1177/1352458515623365 (2016).

107 Gunnarsson, M. *et al.* Axonal damage in relapsing multiple sclerosis is markedly reduced by natalizumab. *Ann Neurol* **69**, 83-89, doi:10.1002/ana.22247 (2011).

108 Dalla Costa, G. *et al.* Serum neurofilaments increase at progressive multifocal leukoencephalopathy onset in natalizumab-treated multiple sclerosis patients. *Ann Neurol* **85**, 606-610, doi:10.1002/ana.25437 (2019).

109 Kuhle, J. *et al.* Fingolimod and CSF neurofilament light chain levels in relapsing-remitting multiple sclerosis. *Neurology* **84**, 1639-1643, doi:10.1212/wnl.0000000000001491 (2015).

110 Piehl, F. *et al.* Plasma neurofilament-light chain levels in patients with MS switching from injectable therapies to fingolimod *Mult Scler* **in press** (2017).

111 Fogdell-Hahn, A. Antidrug Antibodies: B Cell Immunity Against Therapy. *Scand J Immunol* **82**, 184-190, doi:10.1111/sji.12327 (2015).

112 Lundstrom, W. *et al.* Soluble IL7Ralpha potentiates IL-7 bioactivity and promotes autoimmunity. *Proc Natl Acad Sci U S A* **110**, E1761-1770, doi:10.1073/pnas.1222303110 (2013).

113 Gregory, A. P. *et al.* TNF receptor 1 genetic risk mirrors outcome of anti-TNF therapy in multiple sclerosis. *Nature* **488**, 508-511, doi:10.1038/nature11307 (2012).

114 Buhelt, S. *et al.* Relationship between soluble CD25 and gene expression in healthy individuals and patients with multiple sclerosis. *Cytokine* **93**, 15-25, doi:10.1016/j.cyto.2017.04.024 (2017).

115 Morgan, D., Ruscetti, F. & Gallo, R. Selective in vitro growth of T lymphocytes from normal human bone marrows. doi:10.1126/science.181845 (1976).

116 Taniguchi, T. & Minami, Y. THE IL-2/IL-2 RECEPTOR SYSTEM - A CURRENT OVERVIEW. *Cell* **73**, 5-8, doi:10.1016/0092-8674(93)90152-g (1993).

117 Rubin, L. A. *et al.* SOLUBLE INTERLEUKIN-2 RECEPTORS ARE RELEASED FROM ACTIVATED HUMAN LYMPHOID-CELLS INVITRO. *Journal of Immunology* **135**, 3172-3177 (1985).

118 Rubin, L. A., Jay, G. & Nelson, D. L. The released interleukin 2 receptor binds interleukin 2 efficiently. *J Immunol* **137**, 3841-3844 (1986).

119 Rubin, L. A., Galli, F., Greene, W. C., Nelson, D. L. & Jay, G. The molecular basis for the generation of the human soluble interleukin 2 receptor. *Cytokine* **2**, 330-336 (1990).

120 Horiuchi, S. *et al.* Altered interleukin-2 receptor alpha-chain is expressed in human T-cell leukaemia virus type-I-infected T-cell lines and human peripheral blood mononuclear cells of adult T-cell leukaemia patients through an alternative splicing mechanism. *Immunology* **91**, 28-34 (1997).

121 Caruso, C., Candore, G., Cigna, D., Colucci, A. T. & Modica, M. A. Biological significance of soluble IL-2 receptor. *Mediators Inflamm* **2**, 3-21, doi:10.1155/s0962935193000018 (1993).

122 Gaffen, S. L. Signaling domains of the interleukin 2 receptor. *Cytokine* **14**, 63-77, doi:10.1006/cyto.2001.0862 (2001).

123 Keystone, E. C. *et al.* Elevated soluble interleukin-2 receptor levels in the sera and synovial fluids of patients with rheumatoid arthritis. *Arthritis Rheum* **31**, 844-849 (1988).

124 Gotoh, Y. *et al.* Determination of age-related changes in human soluble interleukin 2 receptor in body fluids of normal subjects as a control value against disease states. *Clin Chim Acta* **289**, 89-97 (1999).

125 Sack, U. *et al.* Age-dependent levels of select immunological mediators in sera of healthy children. *Clin Diagn Lab Immunol* **5**, 28-32 (1998).

126 Bruunsgaard, H., Pedersen, A. N., Schroll, M., Skinhoj, P. & Pedersen, B. K. TNF-alpha, leptin, and lymphocyte function in human aging. *Life Sci* **67**, 2721-2731 (2000).

127 International Multiple Sclerosis Genetics, C. *et al.* Genetic risk and a primary role for cell-mediated immune mechanisms in multiple sclerosis. *Nature* **476**, 214-219, doi:10.1038/nature10251 (2011).

128 Beecham, A. H. *et al.* Analysis of immune-related loci identifies 48 new susceptibility variants for multiple sclerosis. *Nature Genetics* **45**, 1353-+, doi:10.1038/ng.2770 (2013).

129 Cerosaletti, K. *et al.* Multiple Autoimmune-Associated Variants Confer Decreased IL-2R Signaling in CD4(+)CD25(hi) T Cells of Type 1 Diabetic and Multiple Sclerosis Patients. *Plos One* **8**, doi:10.1371/journal.pone.0083811 (2013).

130 Scheller, J., Chalaris, A., Schmidt-Arras, D. & Rose-John, S. The pro- and anti-inflammatory properties of the cytokine interleukin-6. *Biochimica Et Biophysica Acta-Molecular Cell Research* **1813**, 878-888, doi:10.1016/j.bbamcr.2011.01.034 (2011).

131 Castell, J. V. *et al.* Interleukin-6 is the major regulator of acute phase protein synthesis in adult human hepatocytes. *FEBS Lett* **242**, 237-239 (1989).

132 Hibi, M. *et al.* Molecular cloning and expression of an IL-6 signal transducer, gp130. *Cell* **63**, 1149-1157 (1990).

133 Heinrich, P. C. *et al.* Principles of interleukin (IL)-6-type cytokine signalling and its regulation. *Biochem J* **374**, 1-20, doi:10.1042/bj20030407 (2003).

134 Taga, T., Kawanishi, Y., Hardy, R. R., Hirano, T. & Kishimoto, T. Receptors for B cell stimulatory factor 2. Quantitation, specificity, distribution, and regulation of their expression. *J Exp Med* **166**, 967-981 (1987).

135 Taga, T. *et al.* Interleukin-6 triggers the association of its receptor with a possible signal transducer, gp130. *Cell* **58**, 573-581 (1989).

136 Rose-John, S. & Heinrich, P. C. Soluble receptors for cytokines and growth factors: generation and biological function. *Biochem J* **300 (Pt 2)**, 281-290 (1994).

137 Lust, J. A. *et al.* Isolation of an mRNA encoding a soluble form of the human interleukin-6 receptor. *Cytokine* **4**, 96-100 (1992).

138 Briso, E. M., Dienz, O. & Rincon, M. Cutting edge: soluble IL-6R is produced by IL-6R ectodomain shedding in activated CD4 T cells. *J Immunol* **180**, 7102-7106 (2008).

139 Narazaki, M. *et al.* Soluble forms of the interleukin-6 signal-transducing receptor component gp130 in human serum possessing a potential to inhibit signals through membrane-anchored gp130. *Blood* **82**, 1120-1126 (1993).

140 Calabrese, L. H. & Rose-John, S. IL-6 biology: implications for clinical targeting in rheumatic disease. *Nat Rev Rheumatol* **10**, 720-727, doi:10.1038/nrrheum.2014.127 (2014).

141 Dominitzki, S. *et al.* Cutting edge: trans-signaling via the soluble IL-6R abrogates the induction of FoxP3 in naive CD4+CD25 T cells. *J Immunol* **179**, 2041-2045 (2007).

142 Acosta-Rodriguez, E. V., Napolitani, G., Lanzavecchia, A. & Sallusto, F. Interleukins 1beta and 6 but not transforming growth factor-beta are essential for the differentiation of interleukin 17-producing human T helper cells. *Nat Immunol* **8**, 942-949, doi:10.1038/ni1496 (2007).

143 Bettelli, E. *et al.* Reciprocal developmental pathways for the generation of pathogenic effector TH17 and regulatory T cells. *Nature* **441**, 235-238, doi:10.1038/nature04753 (2006).

144 Kotake, S. *et al.* Interleukin-6 and soluble interleukin-6 receptors in the synovial fluids from rheumatoid arthritis patients are responsible for osteoclast-like cell formation. *J Bone Miner Res* **11**, 88-95, doi:10.1002/jbmr.5650110113 (1996).

145 Desgeorges, A. *et al.* Concentrations and origins of soluble interleukin 6 receptor-alpha in serum and synovial fluid. *J Rheumatol* **24**, 1510-1516 (1997).

146 Maimone, D., Guazzi, G. C. & Annunziata, P. IL-6 detection in multiple sclerosis brain. *J Neurol Sci* **146**, 59-65 (1997).

147 Navikas, V. *et al.* Increased interleukin-6 mRNA expression in blood and cerebrospinal fluid mononuclear cells in multiple sclerosis. *J Neuroimmunol* **64**, 63-69 (1996).

148 Mendel, I., Katz, A., Kozak, N., Ben-Nun, A. & Revel, M. Interleukin-6 functions in autoimmune encephalomyelitis: a study in gene-targeted mice. *Eur J Immunol* **28**, 1727-1737, doi:10.1002/(sici)1521-4141(199805)28:05<1727::aid-immu1727>3.0.co;2-# (1998).

149 Frei, K., Fredrikson, S., Fontana, A. & Link, H. Interleukin-6 is elevated in plasma in multiple sclerosis. *J Neuroimmunol* **31**, 147-153 (1991).

150 Padberg, F. *et al.* CSF and serum levels of soluble interleukin-6 receptors (sIL-6R and sgp130), but not of interleukin-6 are altered in multiple sclerosis. *J Neuroimmunol* **99**, 218-223 (1999).

151 Yoshizaki, K. *et al.* Isolation and characterization of B cell differentiation factor (BCDF) secreted from a human B lymphoblastoid cell line. *J Immunol* **132**, 2948-2954 (1984).

152 Barr, T. A. *et al.* B cell depletion therapy ameliorates autoimmune disease through ablation of IL-6-producing B cells. *J Exp Med* **209**, 1001-1010, doi:10.1084/jem.20111675 (2012).

153 Lampropoulou, V. *et al.* TLR-activated B cells suppress T cell-mediated autoimmunity. *J Immunol* **180**, 4763-4773 (2008).

154 Schneider, A. *et al.* In active relapsing-remitting multiple sclerosis, effector T cell resistance to adaptive T(regs) involves IL-6-mediated signaling. *Sci Transl Med* **5**, 170ra115, doi:10.1126/scitranslmed.3004970 (2013).

155 Serada, S. *et al.* IL-6 blockade inhibits the induction of myelin antigen-specific Th17 cells and Th1 cells in experimental autoimmune encephalomyelitis. *Proc Natl Acad Sci U S A* **105**, 9041-9046, doi:10.1073/pnas.0802218105 (2008).

156 Hong, C., Luckey, M. A. & Park, J. H. Intrathymic IL-7: the where, when, and why of IL-7 signaling during T cell development. *Semin Immunol* **24**, 151-158, doi:10.1016/j.smim.2012.02.002 (2012).

157 Mackall, C. L., Fry, T. J. & Gress, R. E. Harnessing the biology of IL-7 for therapeutic application. *Nat Rev Immunol* **11**, 330-342, doi:10.1038/nri2970 (2011).

158 Kim, G. Y., Hong, C. & Park, J. H. Seeing Is Believing: Illuminating the Source of In Vivo Interleukin-7. *Immune Netw* **11**, 1-10, doi:10.4110/in.2011.11.1.1 (2011).

159 Jiang, Q. *et al.* Cell biology of IL-7, a key lymphotrophin. *Cytokine Growth Factor Rev* **16**, 513-533, doi:10.1016/j.cytogfr.2005.05.004 (2005).

160 Goodwin, R. G. *et al.* Cloning of the human and murine interleukin-7 receptors: demonstration of a soluble form and homology to a new receptor superfamily. *Cell* **60**, 941-951 (1990).

161 Pleiman, C. M. *et al.* Organization of the murine and human interleukin-7 receptor genes: two mRNAs generated by differential splicing and presence of a type I-interferon-inducible promoter. *Mol Cell Biol* **11**, 3052-3059 (1991).

162 Demaret, J. *et al.* Elevated plasmatic level of soluble IL-7 receptor is associated with increased mortality in septic shock patients. *Intensive Care Med* **40**, 1089-1096, doi:10.1007/s00134-014-3346-0 (2014).

163 Fry, T. J. *et al.* A potential role for interleukin-7 in T-cell homeostasis. *Blood* **97**, 2983-2990 (2001).

164 Crawley, A. M., Faucher, S. & Angel, J. B. Soluble IL-7R alpha (sCD127) inhibits IL-7 activity and is increased in HIV infection. *J Immunol* **184**, 4679-4687, doi:10.4049/jimmunol.0903758 (2010).

165 Monti, P. & Bonifacio, E. Interleukin-7 and type 1 diabetes. *Curr Diab Rep* **14**, 518, doi:10.1007/s11892-014-0518-9 (2014).

166 Traggiai, E. *et al.* IL-7-enhanced T-cell response to myelin proteins in multiple sclerosis. *Journal of Neuroimmunology* **121**, 111-119, doi:10.1016/s0165-5728(01)00433-7 (2001).

167 Liu, W. *et al.* CD127 expression inversely correlates with FoxP3 and suppressive function of human CD4+ T reg cells. *J Exp Med* **203**, 1701-1711, doi:10.1084/jem.20060772 (2006).

168 Gregory, S. G. *et al.* Interleukin 7 receptor alpha chain (IL7R) shows allelic and functional association with multiple sclerosis. *Nat Genet* **39**, 1083-1091, doi:10.1038/ng2103 (2007).

169 Keller, E. T., Fu, Z. & Brennan, M. The role of Raf kinase inhibitor protein (RKIP) in health and disease. *Biochemical Pharmacology* **68**, 1049-1053, doi:https://doi.org/10.1016/j.bcp.2004.04.024 (2004).

170 Ojika, K. *et al.* Hippocampal cholinergic neurostimulating peptides (HCNP). *Prog Neurobiol* **60**, 37-83 (2000).

171 Maki, M. *et al.* Decreased expression of hippocampal cholinergic neurostimulating peptide precursor protein mRNA in the hippocampus in Alzheimer disease. *J Neuropathol Exp Neurol* **61**, 176-185 (2002).

172 Bassi, S., Seney, M. L., Argibay, P. & Sibille, E. Elevated Hippocampal Cholinergic Neurostimulating Peptide precursor protein (HCNP-pp) mRNA in the amygdala in major depression. *J Psychiatr Res* **63**, 105-116, doi:10.1016/j.jpsychires.2015.02.006 (2015).

173 Yeung, K. *et al.* Suppression of Raf-1 kinase activity and MAP kinase signalling by RKIP. *Nature* **401**, 173-177, doi:10.1038/43686 (1999).

174 Farooqi, A. A., Li, Y. & Sarkar, F. H. The biological complexity of RKIP signaling in human cancers. *Exp Mol Med* **47**, e185, doi:10.1038/emm.2015.70 (2015).

175 Saito, Y., Mashimo, M., Nobeyama, A., Murakami, K. & Fujii, T. Hippocampal Cholinergic Neurostimulating Peptide Suppresses Acetylcholine Synthesis in T Lymphocytes. *Biol Pharm Bull* **41**, 1611-1614, doi:10.1248/bpb.b18-00287 (2018).

176 Fujii, T. *et al.* Expression and Function of the Cholinergic System in Immune Cells. *Frontiers in Immunology* **8**, doi:10.3389/fimmu.2017.01085 (2017).

177 Hedstrom, A. K., Hillert, J., Olsson, T. & Alfredsson, L. Alcohol as a modifiable lifestyle factor affecting multiple sclerosis risk. *JAMA Neurol* **71**, 300-305, doi:10.1001/jamaneurol.2013.5858 (2014).

178 Kavaliunas, A. *et al.* Importance of early treatment initiation in the clinical course of multiple sclerosis. *Mult Scler* **23**, 1233-1240, doi:10.1177/1352458516675039 (2017).

179 Holmen, C. *et al.* A Swedish national post-marketing surveillance study of natalizumab treatment in multiple sclerosis. *Mult Scler* **17**, 708-719, doi:10.1177/1352458510394701 (2011).

180 Frisell, T. *et al.* Comparative analysis of first-year fingolimod and natalizumab drug discontinuation among Swedish patients with multiple sclerosis. *Mult Scler* **22**, 85-93, doi:10.1177/1352458515579216 (2016).

181 Hillert, J. & Stawiarz, L. The Swedish MS registry – clinical support tool and scientific resource. *Acta Neurol Scand* **132**, 11-19, doi:10.1111/ane.12425 (2015).

182 Van der Auwera, G. A. *et al.* From FastQ data to high confidence variant calls: the Genome Analysis Toolkit best practices pipeline. *Curr Protoc Bioinformatics* **43**, 11.10.11-33, doi:10.1002/0471250953.bi1110s43 (2013).

183 Packer, J. S. *et al.* CLAMMS: a scalable algorithm for calling common and rare copy number variants from exome sequencing data. *Bioinformatics* **32**, 133-135, doi:10.1093/bioinformatics/btv547 (2016).

184 Wang, K., Li, M. & Hakonarson, H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res* **38**, e164, doi:10.1093/nar/gkq603 (2010).

185 Landrum, M. J. *et al.* ClinVar: public archive of relationships among sequence variation and human phenotype. *Nucleic Acids Res* **42**, D980-985, doi:10.1093/nar/gkt1113 (2014).

186 Scionti, F., Di Martino, T. M., Pensabene, L., Bruni, V. & Concolino, D. The Cytoscan HD Array in the Diagnosis of Neurodevelopmental Disorders. *High-Throughput* **7**, doi:10.3390/ht7030028 (2018).

187 Bolotin, D. A. *et al.* MiXCR: software for comprehensive adaptive immunity profiling. *Nat Methods* **12**, 380-381, doi:10.1038/nmeth.3364 (2015).

188 Shugay, M. *et al.* VDJtools: Unifying Post-analysis of T Cell Receptor Repertoires. *PLoS Comput Biol* **11**, e1004503, doi:10.1371/journal.pcbi.1004503 (2015).

189 P De Jager , I. M. S. G. C. The Genomic map of multiple sclerosis: over 45 novel susceptibility variants and translation of genetics to biology. *Mult Scler* **20**, 497-500, doi:10.1177/1352458514547847 (2014).

190 Uhlen, M. *et al.* A human protein atlas for normal and cancer tissues based on antibody proteomics. *Mol Cell Proteomics* **4**, 1920-1932, doi:10.1074/mcp.M500279-MCP200 (2005).

191 Drobin, K., Nilsson, P. & Schwenk, J. M. Highly multiplexed antibody suspension bead arrays for plasma protein profiling. *Methods Mol Biol* **1023**, 137-145, doi:10.1007/978-1-4614-7209-4_8 (2013).

192 Darmanis, S. *et al.* Identification of candidate serum proteins for classifying well-differentiated small intestinal neuroendocrine tumors. *PLoS One* **8**, e81712, doi:10.1371/journal.pone.0081712 (2013).

193 Kang, H. *Efficient and Parallelizable Association Container Toolbox (EPACTS)*, 2014).

194 Zhan, X., Hu, Y., Li, B., Abecasis, G. R. & Liu, D. J. RVTESTS: an efficient and comprehensive tool for rare variant association analysis using sequence data. *Bioinformatics* **32**, 1423-1426, doi:10.1093/bioinformatics/btw079 (2016).

195 Lee, S., Abecasis, G. R., Boehnke, M. & Lin, X. Rare-variant association analysis: study designs and statistical tests. *Am J Hum Genet* **95**, 5-23, doi:10.1016/j.ajhg.2014.06.009 (2014).

196 Li, B. & Leal, S. M. Methods for detecting associations with rare variants for common diseases: application to analysis of sequence data. *Am J Hum Genet* **83**, 311-321, doi:10.1016/j.ajhg.2008.06.024 (2008).

197 Lee, S., Wu, M. C. & Lin, X. Optimal tests for rare variant effects in sequencing association studies. *Biostatistics* **13**, 762-775, doi:10.1093/biostatistics/kxs014 (2012).

198 lmerTest: Tests for random and fixed effects for linear mixed effect models (Comprehensive R Archive Network (CRAN), 2016).

199 Pollard, K. S., Dudoit, S. & van der Laan, M. J. *Multiple Testing Procedures: R multtest Package and Applications to Genomics, in Bioinformatics and Computational Biology Solutions Using R and Bioconductor*. (Springer, 2005).

200 Sobel, R. A. & Ahmed, A. S. White matter extracellular matrix chondroitin sulfate/dermatan sulfate proteoglycans in multiple sclerosis. *J Neuropathol Exp Neurol* **60**, 1198-1207, doi:10.1093/jnen/60.12.1198 (2001).

201 Papadakis, K. A. *et al.* The role of thymus-expressed chemokine and its receptor CCR9 on lymphocytes in the regional specialization of the mucosal immune system. *J Immunol* **165**, 5069-5076 (2000).

202 Wendt, E. & Keshav, S. CCR9 antagonism: potential in the treatment of Inflammatory Bowel Disease. *Clin Exp Gastroenterol* **8**, 119-130, doi:10.2147/CEG.S48305 (2015).

203 Goldstein, B. A., Hubbard, A. E., Cutler, A. & Barcellos, L. F. An application of Random Forests to a genome-wide association dataset: methodological considerations & new findings. *BMC Genet* **11**, 49, doi:10.1186/1471-2156-11-49 (2010).

204 Gu, B. J. *et al.* A rare P2X7 variant Arg307Gln with absent pore formation function protects against neuroinflammation in multiple sclerosis. *Hum Mol Genet* **24**, 5644-5654, doi:10.1093/hmg/ddv278 (2015).

205 chris.cotsapas@yale.edu, I. M. S. G. C. E. a. & Consortium, I. M. S. G. Low-Frequency and Rare-Coding Variation Contributes to Multiple Sclerosis Risk. *Cell* **175**, 1679-1687.e1677, doi:10.1016/j.cell.2018.09.049 (2018).

206 Jung, S. H. *et al.* Genome-wide copy number variation analysis identifies deletion variants associated with ankylosing spondylitis. *Arthritis Rheumatol* **66**, 2103-2112, doi:10.1002/art.38650 (2014).

207 Wider, C. *et al.* Phactr2 and Parkinson's disease. *Neurosci Lett* **453**, 9-11, doi:10.1016/j.neulet.2009.02.009 (2009).

208 chris.cotsapas@yale.edu, I. M. S. G. C. E. a. & Consortium, I. M. S. G. Low-Frequency and Rare-Coding Variation Contributes to Multiple Sclerosis Risk. *Cell* **178**, 262, doi:10.1016/j.cell.2019.06.016 (2019).

209 Wucherpfennig, K. W. *et al.* Clonal expansion and persistence of human T cells specific for an immunodominant myelin basic protein peptide. *J Immunol* **152**, 5581-5592 (1994).

210 Sethi, D. K., Gordo, S., Schubert, D. A. & Wucherpfennig, K. W. Crossreactivity of a human autoimmune TCR is dominated by a single TCR loop. *Nat Commun* **4**, 2623, doi:10.1038/ncomms3623 (2013).

211 Jun, L., Balboni, A. L., Laitman, J. T. & Bergemann, A. D. Isolation of DNTNP, which encodes a potential nuclear protein that is expressed in the developing, dorsal neural tube. *Dev Dyn* **224**, 116-123, doi:10.1002/dvdy.10090 (2002).

212 Christopher, M. A., Kyle, S. M. & Katz, D. J. Neuroepigenetic mechanisms in disease. *Epigenetics Chromatin* **10**, 47, doi:10.1186/s13072-017-0150-4 (2017).

213 Gris, D. *et al.* NLRP3 plays a critical role in the development of experimental autoimmune encephalomyelitis by mediating Th1 and Th17 responses. *J Immunol* **185**, 974-981, doi:10.4049/jimmunol.0904145 (2010).

214 Pennisi, M. *et al.* Inflammasomes, hormesis, and antioxidants in neuroinflammation: Role of NRLP3 in Alzheimer disease. *J Neurosci Res* **95**, 1360-1372, doi:10.1002/jnr.23986 (2017).

215 Kidd, J. M. *et al.* Mapping and sequencing of structural variation from eight human genomes. *Nature* **453**, 56-64, doi:10.1038/nature06862 (2008).

216 de Cid, R. *et al.* Deletion of the late cornified envelope LCE3B and LCE3C genes as a susceptibility factor for psoriasis. *Nat Genet* **41**, 211-215, doi:10.1038/ng.313 (2009).

217 Niehues, H. *et al.* Late cornified envelope (LCE) proteins: distinct expression patterns of LCE2 and LCE3 members suggest nonredundant roles in human epidermis and other epithelia. *Br J Dermatol* **174**, 795-802, doi:10.1111/bjd.14284 (2016).

218 Jia, X. *et al.* Genome sequencing uncovers phenocopies in primary progressive multiple sclerosis. *Ann Neurol* **84**, 51-63, doi:10.1002/ana.25263 (2018).

219 Wang, Z. *et al.* Nuclear Receptor NR1H3 in Familial Multiple Sclerosis. *Neuron* **90**, 948-954, doi:10.1016/j.neuron.2016.04.039 (2016).

220 cotsapas@broadinstitute.org, I. M. S. G. C. E. a. & Consortium, I. M. S. G. NR1H3 p.Arg415Gln Is Not Associated to Multiple Sclerosis Risk. *Neuron* **92**, 333-335, doi:10.1016/j.neuron.2016.09.052 (2016).

221 Ramagopalan, S. V. *et al.* Rare variants in the CYP27B1 gene are associated with multiple sclerosis. *Ann Neurol* **70**, 881-886, doi:10.1002/ana.22678 (2011).

222 Sato, S. *et al.* Copy number variations in multiple sclerosis and neuromyelitis optica. *Ann Neurol*, doi:10.1002/ana.24511 (2015).

223 Walsh, K. M., Bracken, M. B., Murk, W. K., Hoh, J. & Dewan, A. T. Association between reduced copy-number at T-cell receptor gamma (TCRgamma) and childhood allergic asthma: A possible role for somatic mosaicism. *Mutat Res* **690**, 89-94, doi:10.1016/j.mrfmmm.2010.05.005 (2010).

224 Singh, A. K. *et al.* High Interferon-γ Uniquely in Vδ1 T Cells Correlates with Markers of Inflammation and Axonal Damage in Early Multiple Sclerosis. *Front Immunol* **8**, 260, doi:10.3389/fimmu.2017.00260 (2017).

225 Dudley, E. C., Girardi, M., Owen, M. J. & Hayday, A. C. Alpha beta and gamma delta T cells can share a late common precursor. *Curr Biol* **5**, 659-669, doi:10.1016/s0960-9822(95)00131-x (1995).

226 Margolis, D., Yassai, M., Hletko, A., McOlash, L. & Gorski, J. Concurrent or sequential delta and beta TCR gene rearrangement during thymocyte development: individual thymi follow distinct pathways. *J Immunol* **159**, 529-533 (1997).

227 Muraro, P. A. *et al.* T cell repertoire following autologous stem cell transplantation for multiple sclerosis. *The Journal of Clinical Investigation* **124**, 1168-1172, doi:10.1172/JCI71691 (2014).

228 de Paula Alves Sousa, A. *et al.* Intrathecal T-cell clonal expansions in patients with multiple sclerosis. *Ann Clin Transl Neurol* **3**, 422-433, doi:10.1002/acn3.310 (2016).

229 Muraro, P. A. *et al.* Clonotypic analysis of cerebrospinal fluid T cells during disease exacerbation and remission in a patient with multiple sclerosis. *J Neuroimmunol* **171**, 177-183, doi:10.1016/j.jneuroim.2005.10.002 (2006).

230 Warnke, C. *et al.* Natalizumab affects the T-cell receptor repertoire in patients with multiple sclerosis. *Neurology* **81**, 1400-1408, doi:10.1212/WNL.0b013e3182a84101 (2013).

231 Alfredsson, L. & Olsson, T. Lifestyle and Environmental Factors in Multiple Sclerosis. *Cold Spring Harb Perspect Med* **9**, doi:10.1101/cshperspect.a028944 (2019).

232 Pender, M. P. & Burrows, S. R. Epstein-Barr virus and multiple sclerosis: potential opportunities for immunotherapy. *Clin Transl Immunology* **3**, e27, doi:10.1038/cti.2014.25 (2014).

233 Lundstrom, W. *et al.* Interferon beta treatment of multiple sclerosis increases serum interleukin-7. *Mult Scler* **20**, 1727-1736, doi:10.1177/1352458514532700 (2014).

234 Maier, L. M. *et al.* Soluble IL-2RA Levels in Multiple Sclerosis Subjects and the Effect of Soluble IL-2RA on Immune Responses. *Journal of Immunology* **182**, 1541-1547 (2009).

235 Iaffaldano, P. *et al.* Fingolimod versus interferon beta/glatiramer acetate after natalizumab suspension in multiple sclerosis. *Brain* **138**, 3275-3286, doi:10.1093/brain/awv260 (2015).

236 Frisullo, G. *et al.* CD4+T-bet+, CD4+pSTAT3+ and CD8+T-bet+ T cells accumulate in peripheral blood during NZB treatment. *Mult Scler* **17**, 556-566, doi:10.1177/1352458510392263 (2011).

237 Treumer, F., Zhu, K., Glaser, R. & Mrowietz, U. Dimethylfumarate is a potent inducer of apoptosis in human T cells. *J Invest Dermatol* **121**, 1383-1388, doi:10.1111/j.1523-1747.2003.12605.x (2003).

238 Boulton, C., Meiser, K., David, O. J. & Schmouder, R. Pharmacodynamic effects of steady-state fingolimod on antibody response in healthy volunteers: a 4-week, randomized, placebo-controlled, parallel-group, multiple-dose study. *J Clin Pharmacol* **52**, 1879-1890, doi:10.1177/0091270011427908 (2012).

239 Hait, N. C. *et al.* Active, phosphorylated fingolimod inhibits histone deacetylases and facilitates fear extinction memory. *Nat Neurosci* **17**, 971-980, doi:10.1038/nn.3728 (2014).

240 Mehling, M. *et al.* Th17 central memory T cells are reduced by FTY720 in patients with multiple sclerosis. *Neurology* **75**, 403-410, doi:10.1212/WNL.0b013e3181ebdd64 (2010).

241 Jones, S. A., Scheller, J. & Rose-John, S. Therapeutic strategies for the clinical blockade of IL-6/gp130 signaling. *The Journal of clinical investigation* **121**, 3375-3383, doi:10.1172/JCI57158 (2011).

242 Cohan, S. Therapeutic efficacy of monthly subcutaneous injection of daclizumab in relapsing multiple sclerosis. *Biologics* **10**, 119-138, doi:10.2147/BTT.S89218 (2016).

243 Häggmark, A. *et al.* Plasma profiling reveals three proteins associated to amyotrophic lateral sclerosis. *Ann Clin Transl Neurol* **1**, 544-553, doi:10.1002/acn3.83 (2014).

244 Bystrom, S. *et al.* Affinity proteomic profiling of plasma, cerebrospinal fluid, and brain tissue within multiple sclerosis. *J Proteome Res* **13**, 4607-4619, doi:10.1021/pr500609e (2014).

245 Ayoglu, B. *et al.* Affinity proteomics within rare diseases: a BIO-NMD study for blood biomarkers of muscular dystrophies. *EMBO Mol Med* **6**, 918-936, doi:10.15252/emmm.201303724 (2014).

246 Neiman, M. *et al.* Plasma profiling reveals human fibulin-1 as candidate marker for renal impairment. *J Proteome Res* **10**, 4925-4934, doi:10.1021/pr200286c (2011).

247 Haggmark, A. *et al.* Antibody-based profiling of cerebrospinal fluid within multiple sclerosis. *Proteomics* **13**, 2256-2267, doi:10.1002/pmic.201200580 (2013).

248 Schwenk, J. M., Gry, M., Rimini, R., Uhlén, M. & Nilsson, P. Antibody suspension bead arrays within serum proteomics. *J Proteome Res* **7**, 3168-3179, doi:10.1021/pr700890b (2008).

249 Tsugu, Y. *et al.* High levels of hippocampal cholinergic neurostimulating peptide (HCNP) in the CSF of some patients with Alzheimer's disease. *Eur J Neurol* **5**, 561-569 (1998).

250 Lassmann, H., van Horssen, J. & Mahad, D. Progressive multiple sclerosis: pathology and pathogenesis. *Nat Rev Neurol* **8**, 647-656 (2012).

251 Cree, B. A. C. in *Handbook of Clinical Neurology* Vol. Volume 122 (ed S. Goodin Douglas) 211-230 (Elsevier, 2014).

252 Van Horebeek, L. *et al.* A robust pipeline with high replication rate for detection of somatic variants in the adaptive immune system as a source of common genetic variation in autoimmune disease. *Hum Mol Genet* **28**, 1369-1380, doi:10.1093/hmg/ddy425 (2019).

253 Valori, M. *et al.* A novel class of somatic mutations in blood detected preferentially in CD8+ cells. *Clin Immunol* **175**, 75-81, doi:10.1016/j.clim.2016.11.018 (2017).